

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 14:57:02 ; Search time 534.27 Seconds
(without alignments)
12732.167 Million cell updates/sec

1732.167 Million cell updates/sec

Title: US-08-153-397a-1
3962
1 CGCGCCGGAAGTGGGGTGA.....AAAAAAAAAACCGGAATTC 3962

Sequence: 1 CGCGCCGGAAGTGGGGTGA.....AAAAAAAAAACCGGAATTC 3962

Scoring table: IDENTITY_NUC 1
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

- 1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
- 23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3962	100.0	3962	18	AAAT93785
2	3960.4	100.0	3962	16	AAQ92520
3	3960.4	100.0	3962	16	AAQ92522
4	3617.6	91.3	3754	16	AAQ84782
5	3383.2	85.4	3554	24	AA516842
6	2783.6	70.3	2861	21	AAQ99051
7	2783.6	70.3	2861	22	AAH33198
8	642	16.2	3096	19	AAV48292
9	642	16.2	3157	16	AAQ92521

10	642	16.2	3157	16	AAQ92523	Human colonic aden
11	642	16.2	3157	18	AAAT93784	CCR-2, a human mam
12	639.8	16.1	3120	19	AAV55895	Receptor protein t
13	639.8	16.1	3120	20	AAV55895	Receptor protein t
14	544.2	13.7	2725	22	AAH99505	Human protein enco
15	522.8	13.2	563	23	AA557829	CDNA #505 encoding
16	453.4	11.4	2128	15	AAQ64158	Partial coding seq
17	324.4	8.2	408	21	AAQ00624	Human secreted pro
18	236	6.0	272	16	AAAT2477	Human gene signatu
19	228.4	5.8	463	22	ABA43008	Human breast cell
20	228.4	5.8	463	22	ABA53422	Human foetal liver
21	228.4	5.8	463	22	ABA23197	Human foetal liver
22	228.4	5.8	463	22	AAK01693	Probe #1663 for ge
23	228.4	5.8	463	22	AAK27141	Human brain expres
24	228.4	5.8	463	22	AA111731	Human bone marrow
25	228.4	5.8	463	22	AA133038	Probe #1664 for ge
26	228.4	5.8	463	22	AA101659	Probe #1724 used t
27	182.2	4.6	2820	16	AAAT51456	Probe #1650 used t
28	180.6	4.6	1746	23	AA584464	gd.trkA fusion use
29	180.6	4.6	2301	19	AAV20445	DNA encoding novel
30	180.6	4.6	2301	21	AAQ09303	Human c-trk oncoge
31	180.6	4.6	3194	16	AAAT51457	Human trk oncogene
32	180.6	4.6	3194	21	AAZ88839	Human trkB recepto
33	180.6	4.6	3707	17	AAQ99277	Human trkB recepto
34	180.6	4.6	2526	15	AAQ69029	Human neurotrophic
35	180.2	4.5	2526	13	AAQ28668	Porcine TrkC gene.
36	178.6	4.5	2526	14	AAQ34581	Encodes adult porc
37	178.6	4.5	2526	16	AAAT51458	trkC clone in pFL1
38	178.6	4.5	2940	16	AAAT51458	gd.trkC fusion use
39	177	4.5	2750	20	AAK87599	DNA coding for neu
40	158	4.0	4092	20	AAV07030	Human receptor tyr
41	151	3.8	175	22	ABA48136	Human breast cell
42	151	3.8	175	22	ABA66013	Human foetal liver
43	151	3.8	175	22	ABA33099	Probe #11565 for g
44	151	3.8	175	22	AAK14437	Human brain expres
45	151	3.8	175	22	AAK40171	Human bone marrow

ALIGNMENTS

RESULT 1
AAAT93785 standard; CDNA; 3962 BP.

AAAT93785; 16-FEB-1998 (first entry)

Human mammary carcinoma kinase 10 (MCK-10) cDNA sequence.
Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
Proliferative disease; cancer; insulin receptor family;
tyrosine kinase neurotrophin receptor; MCK-10 activity;
neurological disorder; aberrant expression; ds.

Homo sapiens.

Key CDS location/Qualifiers
321..307 /*tag= a

US5677144-A.

14-OCT-1997.

08-NOV-1994; 94US-0336343.

16-NOV-1993; 93US-0153397.

(ALVE/) ALVES F H E.

(ULTR/) ULTRICH A.

Db	1681	tgltggaagagagagctgacggttcaactctctctgctctctctggtggaaacatcccaacaa	1740
Qy	1741	ACGGCCAGGTCCTAGAGACCAACCCCGTAGACAGAGACCCGGCCCTGTTGGGAATCCG	1800
Db	1741	acgcccagagctctcagagagcaacccccctgacagagccccggcctctgfygaaatccg	1800
Qy	1801	CCACACTCCGGTCCCTGTTGTTCCCAATTTGGCTGTGCTGCTCTCCAAATTCAGCCTAC	1860
Db	1801	ccaactccgcctccctcgctgctcccaatgagctcgcgtgctctcccaatccagctaac	1860
Qy	1861	GCCCTCTTCTGAGCCACTTACAGCCCGTCCCTCTGAGAGCCCGGGCCCCCACCACCGCCT	1920
Db	1861	gctcctctctgacacttaagcccgctccctctgagagccgggccccccaacccgct	1920
Qy	1921	GGGCGAAACCCACACACCAACCTCTACAGTGGGACCTATATGAGCCCTGAGAGCCAG	1980
Db	1921	gggcccacaaaccacacaaacacacagctacagtggygactatacgagcctgagaagccag	1980
Qy	1981	GGCCCGCGTCTGAGCCCGCCACCTCCCGAAGACGCGCCCCCAATTTATGGGAGCGTAC	2040
Db	1981	ggccccgcgtctctgcccccaactcccccaagagcgcccccaatagccgagagctgaca	2040
Qy	2041	TGTCTTACCTTGCAAGGCGCTCACCGGGGCAACCTATGCTGTGCTGTGACCTCCCCAG	2100
Db	2041	ttcttaccctgacagggcgtaaccggggacacacatagtgctgctgacatgccccag	2100
Qy	2101	GGGCAATCGGGGATGGGCCCCCAGAGTGATTTCCCTGATCTGCACCTCCGTTCAAG	2160
Db	2101	gggcagctcggggagctggcccccaagtggaattcccccgatctcgactccgcttcaag	2160
Qy	2161	AGAAGCTTGGCGAGGCGCAATTGTTGGGAGAGTGACCTGTGTGAGTGCAGACGCCCTCAG	2220
Db	2161	agaagcttgcgggggccagcttggygagtgtaactgtgtagggtgcagacagccccaag	2220
Qy	2221	ATCTGGTCACTCTTGAATTTCCCTTAATGTGCGTAAAGGACACCTTTGCTGTAGCTG	2280
Db	2221	atctggctcagctcttgatcttcccccttaaatgtgtgaaggacaccccttgctgtagctg	2280
Qy	2281	TCAAGATCTTACGCGCAGATGCGACCAATAATGCCAGCTTCTCTGTCTCCAGGAAG	2340
Db	2281	tcaagatcttaagccaggttcgacccaabaatgccaagcttccctgtctcccaagaaag	2340
Qy	2341	ATTTCCTGAAAGGTGAAGATCATGTGAGGCTCAAGGACCCCAATCATTCGGCTGC	2400
Db	2341	attctctgaaagaggtgtaagatcatgtctgaagctcaaggacccaacatcatctgcgtcg	2400
Qy	2401	TGGGGCTGTGTGGCGAGGACGCCCTGTGATATTTCTACATCATGAGAGACGGCG	2460
Db	2401	tggggctgtgtgtgcagagagcaacccccctctgtaatactcgaactcaatgagaaacggtg	2460
Qy	2461	ACCTCAACCAAGTTCCTCAGTGGCCAACAGCTGAGAGACAAAGCAGCCGAGGGGCCCTG	2520
Db	2461	acctcaaccaaagtcttccctcagtgccaacagctgtaggaacaaggccggggggccctg	2520
Qy	2521	GGGAGGGGAGGCTCCGCAAGGGGCCACCATAGCTAACCAATGCTGCTGCAATGGGACG	2580
Db	2521	gggaggggagagctgcgcaaggggcccacatagatcccaatgctgtcgtacatgtgacg	2580
Qy	2581	CCACAGATGCGCTCCGCGATGCGCTATCTGTTGGCCACACTCAACTTGTACATCGGGAC	2640
Db	2581	ccacagatgcctcccgagatgagctatctgagccacactcaacttcaacatcgvggaccgg	2640
Qy	2641	CCACGCGGAAGTGCCTAGTTGGGAAAATTTCCACATCAAAAATCGCAGACTTTGGCATGA	2700
Db	2641	ccaacgggaactgctagctggygaaaatttaccacataaaatcgacagacttggcatga	2700
Qy	2701	GCCGGAACCTCTATCTGTGGGACATTAACCGTGTCCAGGGCCGGGCAATGCTGCCAATCC	2760
Db	2701	gcgcgaactctatgtctggygaactatcacggtgacagggccgggacgtgcgccatcc	2760
Qy	2761	GCTGATGCGCTGGGAGTGCATTCCTCATGGGAAATTTACACACTCGAGTACGTGCTGG	2820

Db	2761	gctcgatggtccctggaagatgacatccatctggaagaaatctcaagactgcgaatgacgtgtgag	2820
QY	2821	ccatttgatgagccctgtagtgagagtgctgaagacgactgtatgagggccacagcccttttgaggacac	2880
Db	2821	cccttggtgtgagccctgtaggaaggtgcgtgacgtctgtatgagccagcccttctggtgacacg	2880
QY	2881	TCACGACGACGACGAGCTCATCGAAGACGCGGGGGAGTTCTTCCGGGACCAAGGCGCGGACAG	2940
Db	2881	tcacacgacgagacaggttcatcagaaacggtgaggtctcttcggaaccaggtgcgcgcagag	2940
QY	2941	TGACACTGTGTCGGGGGGGCGCTGCTGCTCCCGGAGGGCCCTATATAGACTGATGCTTCGCTGCT	3000
Db	2941	tgtaacctgtcccggtgcgcgcctgtgcctgcgcgcagggccatatagtacgtgtcttcgtgtct	3000
QY	3001	GGAGCGGGGAGTCTAGACAGCGACACACCCCTTCTCCAGCTGCAATCGGTTCTGCGGACGAG	3060
Db	3001	ggagccggaggtctgtagacagcgacacaccccttccacagcgatcgtgtctctgtgcagagag	3060
QY	3061	ATGCACTCAACAGGTGTGTAATTCACACATCCAGCTGCCCTCCCTCAGGAGAGTATCACAG	3120
Db	3061	atgcatctcaacacggtgtgtaatcaacaatccagtgctccctccataggagtgtatcag	3120
QY	3121	GGGAAGCCAGTGTACACTAAACACAGAGACACATGCACTTGCTGCCCTTCCCTCCCGA	3180
Db	3121	gggaagccaggtgtacactaaacaaagaaacacatgtcacctctgccttccctccacga	3180
QY	3181	CAGCCCATACCTCTTAATATAGAGGAGAGACGTGACGTGAGGTGGGGCCACCCAGGAG	3240
Db	3181	cagcccatcaactcctcaatataagagacagatcgtcaggtgtgtctggtgccacacagggag	3240
QY	3241	CTGATGCCCTTCTGCTCCCTTCTGAGACACTGTCAATGTCCCTTCTGTTCTTCTTCC	3300
Db	3241	ctgatgcccttctccttcccttctctctctctgacacactcctcatgtccctctgtcttctcc	3300
QY	3301	TGGAAGCCCGTGTGCGCCACCCACCACTGCTGTGTGATGGGATCTCTTCCACCTTCCTCT	3360
Db	3301	tggaaagccctgtctgcgcacacacacagctgtctctgtgtatgtgatactctccacccctct	3360
QY	3361	AGCCATCCCTTGTGGGGAAGGAGTGGGGAAATATAGATATAGACACTGAGACATGGCCATG	3420
Db	3361	agccatcccttgtgggaaggtgtgggaaatatagatatagatacctgtgaaatgagccatctg	3420
QY	3421	GAGCACTGGGCCCCACTGTGACAACTGATTCTGTGAGAGGTGCTGCGCCACAGCTTC	3480
Db	3421	gagcacctgggccccactgtgacaacactgttctctgtgagagtgtgtctgcgcgccacgtctc	3480
QY	3481	TCTCTCCCTGTTCACACACTGTGACCCCCACTGCTGTGAAATCTGGGGGTAGAGAGCAAGA	3540
Db	3481	tctctccctgttcaaacactgtgacccccactgtctgtgaaatactgtgaggtgtgagtgagaa	3540
QY	3541	AGGAAGAGAAATATTCTTCTGTGATCCGCTGATGCTTGTCTCTAGCTTGGGCTTCTTC	3600
Db	3541	aggaagagaaaaatgttctctgt	3600
QY	3601	CTCTCTCATCACTGAAACACTGTGACCTGTGGGGTACCCCGCCACAGCCTCATGTACCC	3660
Db	3601	ctctctcatcaactgtgaataacactgtgactgtgggttagccccgcgccagccctcatgtaccc	3660
QY	3661	CCACTTCCCACTTTCAGTCTTTGTATGACTTGAACCTTCTTAAAGCCTTATACGTTTCTGTGAG	3720
Db	3661	ccaacttcccaactgtcagctctctgtatgtaaaccttctctaagcctatacgttcttctgtgag	3720
QY	3721	TAAATATTGGGATTTGGGGGGAAGAGGAGCAACGGCCCATAGCTTGGGTTGGATCATC	3780
Db	3721	taaatattgggaatttggggggaagaaagaggaacagggccataagctctgtgggtgtgtgaatc	3780
QY	3781	TCTAGTGTAGCTGCGACATGATTTTTCATATATCACTTGGGGTTGTACATTTTGGG	3840
Db	3781	tctagtgtagctgacaaatgtatttctcataatacaacttgggtgtgtgtgtgtgtgtgtgtgtgt	3840
QY	3841	GGAGAGACACAGATTTTATACATTAATATATGACCTTACGCTTGAGGCAATTTTAAATCCCT	3900
Db	3841	ggagagacacagatlttatacaataataatgtgacctcatgtgtgaagcaatttaatccct	3900

```

OY 3901 GCCTAGCAGGTATTAATAAGCTTGTGAGTTTCACCAAAAAAAAAAACCGGAT 3960
DB 3901 gcactagcaggtataataaagcttggtttccacaaaaaaacacggaat 3960
OY 3961 TC 3962
DB 3961 tc 3962

RESULT 2
AAQ92520
ID AAQ92520 standard; cDNA to mRNA; 3962 BP.
XX
XX AAQ92520;
XX
XX 26-NOV-1995 (first entry)
XX
XX Human mammary carcinoma kinase 10 (MCK-10) cDNA.
XX
XX Mammary carcinoma kinase 10; transmembrane receptor;
XX receptor tyrosine kinase; cancer; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 321..3080
XX FT /*tag- a
XX FT misc_difference 2315
XX FT /*tag- b
XX FT /*note- "some clones have 6 AA deletion here"
XX
XX MO9514088-A.
XX
XX 26-MAY-1995.
XX
XX 16-NOV-1994; 94MO-EP03797.
XX
XX 16-NOV-1993; 93US-0153397.
XX
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX Alves FHE, Ullrich A;
XX
XX WP1; 1995-224054/29.
XX P-PSDB; AAR75502.
XX
XX New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
XX derived vectors, transformed cells, proteins and antibodies useful
XX for diagnosis and treatment of proliferative disease, esp. cancer,
XX and for screening modulators
XX
XX Claim 4; Page 50-52; 115pp; English.
XX
XX cDNA prepd. from human breast cancer cell line MCF7 (ATCC HTB22) and
XX used in a PCR with two degenerate oligo primer pools based on
XX conserved sequences of the kinase domain of receptor tyrosine
XX kinases. One clone, designated MCK-10, was identified as novel RTK.
XX The PCR fragment was used to screen a lambda gtl1 library of human
XX fetal brain cDNA. Several overlapping clones were identified. The
XX composite of these cDNA clones is given in AAQ92520 and the deduced AA
XX sequence in AAR75502. Some of the clones had a deletion of 6 AAs at
XX posn. 2315 in the MCK-10 sequence. MCK-10 has all the
XX characteristics of a receptor PK (see AAR75502 FT). Screening of
XX human placental library yielded two cDNA clones MCK-10-1 and
XX MCK-10-2. One of the clones isolated from the human fetal brain
XX library contd. an additional 18 nts in the TK domain. The MCK-10 splice
XX isoforms have been designated MCK-10-1 (with an additional 111 bp
XX between nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3
XX (with the additional 111 bp and and 18 bp in the TK domain); and MCK10-4
XX (with the additional 18 bp). The predicted mol. wts. of MCK-10-1 and
XX MCK-10-2 precursors are 101.13 and 97.17 kD respectively, and can thus
XX be subdivided into a 34.31 kD alpha subunit and and 66.84 or 62.88 kD

```

```

CC beta subunits that contain the TK homology and alternative splice sites.
XX
XX Sequence 3962 BP: 735 A; 1235 C; 1181 G; 811 T; 0 other;
XX
Query Match 100.0%; Score 3960.4; DB 16; Length 3962;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3961; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CGGGCTGAGACTGGGAGCTGGGAGCTAGAGAAATCTGAGTGAGGCGCCGAGAG 60
DB 1 cgggcctgagactgggagctgggagctaggagaaatctgagtgagggcccgagag 60
OY 61 CTGCTCTCGGAGAGCGCCCTCCGACACCCGAGCCCGCGCGCTCCGCTCCGAGCTC 120
DB 61 ctgctctcggagagcgccctccgacacccgagcccgcgcgctccgctcccgagctc 120
OY 121 CCGGCTCCTGCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 180
DB 121 cgggctcctgctccctccgctccgctccgctccgctccgctccgctccgctccgct 180
OY 181 CCGGAGTGGAGGCTGGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
DB 181 cgggagtgagagctgggctgggagagagagagagagagagagagagagagagagat 240
OY 241 TCAGTGAAGCAGTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 300
DB 241 tcagtgaagcagtgaggagtggagtggagtggagtggagtggagtggagtggagtgg 300
OY 301 GGGCCGAGGAGATCAGAGAGCTATGGAGACAGCCCTGTCTATCTTACTGCTGCTCT 360
DB 301 gggccgagagatcagagagctatggagacagccctgtctatcttactgctgctct 360
OY 361 TGGTGGCAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 361 tggtggcagtggagatgatgatgatgatgatgatgatgatgatgatgatgatgatgat 420
OY 421 CCCTGGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 ccctgggagtggagagagagagagagagagagagagagagagagagagagagagag 480
OY 481 CAGATTCACATGCGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 cagattcacatgcgcccgccgacagagagagagagagagagagagagagagagag 540
OY 541 GCCCGCAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
DB 541 gcccgccagagaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 600
OY 601 TCCACTGTGTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 601 tccactgtgtgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
OY 661 TCTCCCGAGAGTACCGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 661 tctcccgagagtaccggctgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
OY 721 ACCCGTGGGAGCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 acccggtggagcagagagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
OY 781 ACCTGGGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 781 acctgggagggagagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
OY 841 TGAATGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 900
DB 841 tgaatgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 900
OY 901 ACACCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 901 acacgccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960

```

```

OY 961 CCTATGACGACATACCGTGGCGGACTGCTAGTATGGGGGTCTGGCCACTGGCAGATG 1020
    |||||
Db 961 cccatgacgacataccgttggcgcgactgcagctatgggtctctggccagctggcagatg 1020
OY 1021 GTGTGGGGGGGTGATGACTTTAGGAAGATGACAGACGCTGGGCTGTGGCCAGCTATG 1080
    |||||
Db 1021 gctgtgttgggctcgcgaagactttagaagaagctcagagcgcgcgtcttgcgcagctatg 1080
OY 1081 ACTATGTGGATGAGACCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGATTGAGT 1140
    |||||
Db 1081 acctatgtgagatgagcaacacacagcttcccgctgtctgttggagatgtgagatctgagc 1140
OY 1141 TTATACCGGCTGAGAGGCTTTCACGCTATGACGATGCTTAAACATGACAGCCTGG 1200
    |||||
Db 1141 ttgacggcgcgcgaagccttccagcctatgagcttccacgtttaaacaatgcaacgcttgc 1200
OY 1201 GAGCCGCTGCTGGCGGGGGGTGGAATGTCGCTTCCGGCGCTGGCCCTGACATGGCTGG 1260
    |||||
Db 1201 gagccgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1260
OY 1261 AGGCGGAGCCCATGCGCACAACCTAGGGGGCAACTGGGGGAGCCGAGGCCGGGCTG 1320
    |||||
Db 1261 agggggagcccatgctgcgcacaacctgagggcaacctggggagcccgagccgcgcgcgcgc 1320
OY 1321 TCTCAGTGGCCCTTGGCGGCGCGCTGCTGCTTCTGCAAGTGCCTTCTTCTTGGCG 1380
    |||||
Db 1321 tctcagtgcccttggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
OY 1381 GGCCTGCTTACCTCTCAGCAGAAATCTCTCATCTCTCATGTGTGAGAACATCTCTTC 1440
    |||||
Db 1381 ggcctcgttactcttctcagagaaatcctctcatctctgagtgtgaaacaatctctctctc 1440
OY 1441 CGGCACTGGGAGCACTTCCCGCAGCCCTCGTGGCGCGCTGCGCTGCGCACCCTCCACA 1500
    |||||
Db 1441 cggcaactgggagagcaacttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1500
OY 1501 ACTTACAGCAGCTTGGAGCTGAGCCAGAGCCAGCAGCCCGTGGCCAAAGCCGAGGGA 1560
    |||||
Db 1501 acttaagcagcttggagcctgagcagccagagccagcagccgtgscagagccgagggga 1560
OY 1561 GCGCCAGCGCATCTCTCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
    |||||
Db 1561 gcccagacgcacatctctcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1620
OY 1621 TTGCCCTATGCTCTGGGGGCTGCACTGGCGCAGGCTCTTCAAGAGGCTGAACGGAGGG 1680
    |||||
Db 1621 ttgccctatgtctcgtgggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1680
OY 1681 TGTGTGAAGAGAGAGCTGACGGTTCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
    |||||
Db 1681 tgttgaagagagagcctgacgcttcaactctctctctctctctctctctctctctctctctca 1740
OY 1741 ACCGGCCAGGCTCTGAGAGCAACCCCGGTACACAGAGACCCCGGCTGCTGGGAATCCGC 1800
    |||||
Db 1741 accggccaggtctctagagagcaccgccgttaccagagaccgcgcgcgcgcgcgcgcgcgcgc 1800
OY 1801 CCACTCCGCTCCCTGTGTCCCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
    |||||
Db 1801 ccaactccgctccctgtgtccccaatgagctgtcgtctcgtctcgcacatccagcctacc 1860
OY 1861 GCTCTCTCTGGCACAATTACGCCGCTCCCTCGAGGCCCGGGCCGCCACACCCGCT 1920
    |||||
Db 1861 gctctctcttctgcaactaagcgcgcgtccctctgagccgcgcgcgcgcgcgcgcgcgcgcgc 1920
OY 1921 GGGGCAAAACGACCAACCAAGCGCTACATGGGAGATATATGAGGCTTGAAGCCAG 1980
    |||||
Db 1921 gggccaaaacccacaacacccagcgtacagcttgggactatatagtagagcgttagagccag 1980
OY 1981 GCGCCCGGCTTCTGCCCCACCTCCCAAGACAGAGCTCCCAATTATGCGGAGGTGACA 2040
    |||||
Db 1981 gcgcccgcttctgccccacactccccaagacagctccccaattatgacgagaggtctgaca 2040
OY 2041 TTGTTACCTGTCAGAGGCGCTCACCGGGGGCAACACCTATGCTGCTGCTGCTGCTGCTG 2100
    |||||

```

Db	3121	gggaagccagtgacactaaacaagygagcacatctgcaacctctgccttcacctccga	3180
QY	3181	CAGCCCATCACCCTCTAATAGAGCACTGAGACTGCAAGTGGGCTGGGCCACCAAGGAG	3240
Db	3181	cagcccatcacctctaataagagcactgagactgcagctgagggctctggccaccagggag	3240
QY	3241	CTGATGCCCCCTTCTCCCTCTCCCTGAGACACCTCATGTCGCCCCCTTCCTGTTCTTCCTCC	3300
Db	3241	ctgatgcccccttctcccttctcctctctgagacacactcatctatctcccttccctgtctcttc	3300
QY	3301	TAGAAAGCCCTGTGCCCCACCCACAGCTGTCGTGTGGATGGGATTCCTCCACCCCTCCTCT	3360
Db	3301	tagaaagccccctgtcgccaccaccagctgctcgtctgtagtgatgtctctccaccacctctct	3360
QY	3361	AGCCATCCCTGGGGGAAAGGAGGTGGGGAATATATAGATATGACATCGACATGGCCCATTG	3420
Db	3361	agccatcccccttggggaagggctggggaataataagagataagacaactgagacatgagcccatg	3420
QY	3421	GAGCACCCTGGGCCCCCACTGAGACAACTATTCTCTGAGAGGCTGCTCGGCCACAGCTTC	3480
Db	3421	gagcaccctgggcccccacttgagacaacactgatctcccgagaggttgctcgccccagcttc	3480
QY	3481	TCTCTCCTCTCACACACTGAGACCCCATCTGCTGAGAAATCTGGGGGTGAGAGACAAAGA	3540
Db	3481	tctctcccttctcacacacttgagaccacactgctctgagaaatctggggttagagagacaaga	3540
QY	3541	AGGAGAGAAATAGTTTCCCTGTGCGCTGCTCTGTATGTCATGTCACAGTTTGGGCTCTTC	3600
Db	3541	aggagagaaatagtttcttccctgtgcctgctcctgtaactgtctcctcaagcttggctcttc	3600
QY	3601	CTCTCCATCACCTGAAACACTGGAAGCTGGGGGTAGCCCGGCCAGCCCTCAGTCAACC	3660
Db	3601	ctctccatcacctgaaacacttgagacctggggtagccccgccccagccctcagtaacc	3660
QY	3661	CCACTTCCCACTTGCAAGTCTTGTAGTATGAATTTCTTAAGCTTAACGTTTCTGTGAG	3720
Db	3661	ccacttcccaacttgcaagtcttgtagatagaacttctctaaagctataagttctctgtgag	3720
QY	3721	TAAATATTTGGATTGGGGGAAAGAGGAGGAGCAAGGCCCATAGCCTTGGGGTTGGACATC	3780
Db	3721	taaatatttggatttgggggaaaggggagcgcccatagccttgggggttggacatc	3780
QY	3781	TCTAGTGTAGCTGCCACATGATTGTTTCTTAATATCACTTGGGGTTGTACTTTTGGGG	3840
Db	3781	tctagtgtagcttgcacatctgatttctcatalcatcacttgggttctgacatcttctggg	3840
QY	3841	GGAGAGACACAGATTTTTCACATAATATATGAGACTAGCTTGAAGGCATTTTATCCCT	3900
Db	3841	ggagagacacagatcttctcactaataataagaccatagcttggagcaattttaatccct	3900
QY	3901	GCACTAGGACGATTAATTAAGGTTTGAATTTCCACAAAAAATAAAAAACCGAAT	3960
Db	3901	gcactagagcagtaataataaagtttgatttccacaaaaaataaaaaaaccggaat	3960
QY	3961	TC 3962	
Db	3961	tc 3962	

RESULT	3
AA092522	
ID	AA092522 standard; cDNA to mRNA; 3962 BP.
XX	
XX	AA092522;
AC	
XX	
DT	26-NOV-1995 (first entry)
XX	
DE	Human mammary carcinoma kinase 10 (MCK-10) cDNA.
XX	
KW	Mammary carcinoma kinase 10; MCK-10; transmembrane receptor
KW	receptor tyrosine kinase; cancer; ss.
XX	
OS	Homo sapiens.

XX		
FH	Key	Location/Qualifiers
FT	CDS	321..3080
FT		/tag= a
FT	misc_difference	2315
FTm		/tag= b
XX		/note= "some clones have AA deletion here"
XX	N09514089-A.	
XX	26-MAY-1995.	
XX	16-NOV-1994;	94WO-EP03799.
XX	16-NOV-1993;	93US-0153397.
XX	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.	
XX	Alves FHE, Ullrich A;	
XX		
DR	WPI: 1995-224055/29.	
DR	P-PDB; AAR75504.	
PT	New nucleic acid encoding CCK-2 receptor tyrosine kinase - and derived vectors, transformed cells, proteins and antibodies, useful for diagnosis and treatment of proliferative and nervous system diseases and for screening modulators	
XX		
DS	Disclosure; Page 67-69; 11pp; English.	

CC cDNA p1cd4 from human breast cancer cell line MCF7 (ATCC HT822) was
CC used in a PCR with two degenerate oligo prime pools based on
CC conserved sequences of the kinase domain of receptor tyrosine
CC kinases. One clone, designated MCK-10, was identified as novel RTK.
CC The PCR fragment was used to screen a lambda gt11 library of human
CC fetal brain cDNA. Several overlapping clones were identified. The
CC composite of these cDNA clones is given in AA092822 and the deduced AA
CC sequence in AA075504. Some of the clones had a deletion of 65A at posn
CC 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of
CC a receptor PTK (see AA075304 FT). Screening of human placental library
CC yielded two cDNA clones. One of the clones isolated from the human
CC fetal brain library contained an additional 18 nts in the TK
CC domain. The MCK-10 splice isoforms have been designated MCK-10-1
CC (with an additional 111 bp between nts 1832 and 1943); MCK-10-2
CC (without any insertions); MCK-10-3 (with the additional 111 bps and
CC 18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).
CC The predicted mol. wts. of MCK-10-1 and MCK-10-2 preceptors are
CC 101.13 and 97.17 kD respectively, and can thus be subdivided into a
CC 36.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that
CC contain the TK homology and alternative splice sites.

Query Match	100.0%	Score 3960.4	DB 16	Length 3962
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 3961	Conservative	0	Mismatches	1
			Indels	0
			Gaps	0

QY	1	CGGGCTCGAGAGCTAGTGGGGTGAATGGAGACTTAAGAAATATCGAGCTGGAGAGGCCCGCCAG	60
Db	1	ccggcccccgagagctcvgggtcgagctcvggagacctcaagaaatctcgagagctcggagagcccccgacag	60
QY	61	CTGCTCTTCGGGAGAGCGCGCTCCGACACCCAGACCCCGCCGCGGCTATCCGCTCCGGGCTC	120
Db	61	ctgctcttcgggagagcgcgtctcccgagaaacacgcgagccccgcgcgctccgctcccgctc	120
QY	121	CGGGCTCGTAGGTCCTCTCGCGCTTCGCCGCGCTCCGCGCCGCGCGAGAGAGCCCGGCT	180
Db	121	ccggctctcgtctctccctcgcgctcccccgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	180
QY	181	CCCGGGCTCGGAGCGCTATGGGTCGTGCGCGGAAAGAGCGATGAGAGGTATCTTAAGGTGCTAT	240
Db	181	cccgagctcggagagcctctvggtgtctcgcgcggagaaagacgaaatgaaagtgctctgaaagctgcatac	240

QY 241 TCACGTAGCATGGGCTTGGACTTGAAGAAATGCCAAGAGATGCTGCCACCCTTA 300
|||||
Db 241 tcaactagcagatgggtgtgaacttgaagaaatgcgaagaaatgtcgcaccaaccctta 300
QY 301 GGCCTGAGGATCAGAGCTATGGGACAGAGGCCCTGTATCTTTACTGCTGCTCT 360
|||||
Db 301 ggcctgagagatcagagctatggaccagagccctgtactcttactgtctgtctc 360
QY 361 TGGTGGCAATGAGATGCTGACATGAGGAGCAATTTGATCCTGCCAAGTCCGCTATG 420
|||||
Db 361 tggltggcaatgagatgctgacatgagagacatttgaatccgtccaagtccgctatg 420
QY 421 CCTGGGATGACAGAGCCGACCATCCACAGATGACATCTGCTCCAGCTCTGCT 480
|||||
Db 421 ccttggatgacagagccgaccatcccaagatgacatctctgtctccagctctgt 480
QY 481 CAGATTCCACTGCGCGCCGACAGCAGGTGAGAGACAGTGAAGGAGATGGGCTGCT 540
|||||
Db 481 cagattccactgcgccgcacacagcaggttggagagcagtgacggggtctgt 540
QY 541 GCCCGCAGGAGGTGGTTCCTCCCAAGAGAGAGAGTACTTGCAGAGTGTACACAG 600
|||||
Db 541 gcccgagaggttgggtctcccaagagagagtagtacttgcaggtgtgactacaagac 600
QY 601 TCCACCTGTGCTCTGTGGGACACCAGAGAGCATGCGGGGGGCTGGCAAGAGT 660
|||||
Db 601 tccacctgtgtctgtgtgggacaccagagagcagtcgcyggggtctgtgcaagagt 660
QY 661 TCTCCGAGACTACCGGCTGCTTACTCCGGAGTGGTCGCGCTGAGTGGGTGAAG 720
|||||
Db 661 tctccgagactaccggctgttactccggagtgtcgcgctgtgagtgtgctggaag 720
QY 721 ACCGCTGGGGTCAAGAGTGTATCTCAGGCAATAGAACCTTGAGGAGTGTCTGAAG 780
|||||
Db 721 accgctggggtcagaggtatctcaggaatgagaccctgagaggtgtgtgtaag 780
QY 781 ACCTTGGGCCCCCATGATGTTGGCCGACTGTTGCTTCTACCCCGGGCTGACCGGCTCA 840
|||||
Db 781 accttgggcccccatgattgtggccgactgttcttctacccgggctgacccggctca 840
QY 841 TGAATGTCTGTCTGCGGGTAGAGCTTATGGCTGCTCTGGAGGAGTAGACTCTGTCTT 900
|||||
Db 841 tgaatgtctgtctgcgggtagagcttattggctgctctggagagtagactctgtctt 900
QY 901 AACCGCCCCCTGTGGGGCAGACATATATTATCTGAGGCCGTGTACCTCAACGACTCA 960
|||||
Db 901 aacccgccccctgtggggcagacatattattcttctgagccgtgtacccaagactca 960
QY 961 CCTATGAGGACATATACGTTGGGGGAGCTGACATAGGGGTTCTGGGCGACGTCAGATG 1020
|||||
Db 961 cctatgagacataacggttggggagctgacataggggtctcggccagctgtgcagatg 1020
QY 1021 GTGTGGTGGGCTGATGACTTTAGAGAGAGTCAAGAGCTGCGGGTCTGGCCAGGTATG 1080
|||||
Db 1021 gtgtgggtggctgatgacttttagagagagtcaagagctgcggtctggccaggtatg 1080
QY 1081 ACTATGTGGGATGAGCAACACAGCTTCTCAAGTGGCTATGTGAGATGGAATTTGAGT 1140
|||||
Db 1081 actatgtgggatgagcaaacacagcttctcagtgctatgtgagatggaatttgaft 1140
QY 1141 TTGACCGGCTGAGGGCTTCCAGGGATGACAGTCCAGTGAACATGACACGCTG 1200
|||||
Db 1141 ttgaccggctgagggcttccaggatgacagttcaggttccacacgtaacacgcaagctg 1200
QY 1201 GAGCCGCTTGCCTGGCGGGGTGAATGTCTGCTCGCGGTGACCTTGCATGACCTGGG 1260
|||||
Db 1201 gagccgcttgcctggcggggtgaatgtctgctcgggtgacccctgcattggtcctg 1260
QY 1261 AGGGGAGCCCATGCGCCCAACCTTAGGGGCAACCTGCGGGGACCCAGACCCGGGCTG 1320
|||||
Db 1261 aggggagcccatgcgcccaaccttagggggaaccttgggggaacccagagccggctg 1320

QY 1321 TCTCAGTCCCTTGGGGCGGTGGCTGGCTTCTGCAATGCGCTTCTCTTTGCGG 1380
|||||
Db 1321 tctcagtcccttggggcggtggctggcttctgcaatgcgcttctcttttgcgg 1380
QY 1381 GGCCTGTTACTTCTCAGGAAATCTCTCATCTCTGATGTGATGAACATTTCTCTC 1440
|||||
Db 1381 ggcctgttacttctcaggaatctctcatctctctgtatgttggtaaacattctctc 1440
QY 1441 CGGCACTGGAGAGCATTCTCCGCAAGCCCCCTGTGGTGGCGGCTGGCCACTCCACCA 1500
|||||
Db 1441 cggcactggagagcattctcccgacagccccctgtgtggcgcttggccacctccacca 1500
QY 1501 ACTTCAGAGCTTGGAGCTGAGGCCAGAGGCCAGAGCCCCGTGGCCAAAGCGGAGGA 1560
|||||
Db 1501 acttcagagcttggagcttggagagccagagcccgcttgcgaaggtccgaaggagga 1560
QY 1561 GCCCGACCGCATCTCATGCGCTGCTGGGACATCTGCTCTGCTGCTATCA 1620
|||||
Db 1561 gcccgacccgcatctcatctatctgtctgtgttggatccatctctctgtctctatca 1620
QY 1621 TTGCTCTATGCTCTGCGGCTGCTGCGCAGGCTCTCTACAGAGCTGAACGAGGG 1680
|||||
Db 1621 ttgctctatgctctgcggctgtgctgctgctgctgctcctcagcaagctgaacggagg 1680
QY 1681 TGTGGAAAGAGAGCTGAGGCTTCACTCTCTGCTCTGGGGACATATCTCATCAACA 1740
|||||
Db 1681 tgttggaaagagagctgagagcttcaactctctctctcttgggacatctcatcaaca 1740
QY 1741 ACCGCCAGTCTCTAGAGACCAACCCCTACACAGAGAGCCCGGCTCTGGGAAATCCGC 1800
|||||
Db 1741 accgccagttctctagagagcaaccccgtacacagagcccgctctgtggaaatccgc 1800
QY 1801 CCGACTCGCTCCCTGTGTCCCAATGGCTCTGCTGCTCTCAATCAAGCTTAC 1860
|||||
Db 1801 ccgactcgctccctgtgtcccaatggctctgctgctctcaatcagcttacc 1860
QY 1861 GCTCTCTTGGGACATTAAGCCCGTCCCTCTGAGAGCCCGGGCCCCCACCACCGGCT 1920
|||||
Db 1861 gctctcttgggacattaaagcccgctccctctgagagccgggccccccacacccgct 1920
QY 1921 GGGCCAAACCCCAACACACCCAGGCTTACAGTGGGACTTATGAGAGCTGAGAACCA 1980
|||||
Db 1921 gggccaaaccccaacacacccaggtctacagttggaactatagagccttgaagagcg 1980
QY 1981 GCGCCCCCTTGTGCCCCACCTCCACAGAACGCTCCCATTAATGCCAGAGCTGACA 2040
|||||
Db 1981 gcgcccccttgtgcccccacctccacagaaagcttcccccatatgcccaggtcagaa 2040
QY 2041 TTGTTACCTGAGGAGGCTTACCGGGGCAACCTATGCTGTGCTGCTGCTGCCAC 2100
|||||
Db 2041 ttgttacctgagaggttaccgggtacggtggggcaacacatctgtgcttgcagctccc 2100
QY 2101 GGGCAGTGGGGATGGGCCCCCAGAGTGGATTTCCCTGATCTGACATCCGTTCAAG 2160
|||||
Db 2101 gggcagtggggttggcccccaagagtgtatcttccctgatctgacatccgcttcaag 2160
QY 2161 AGAAGTTGGCGAGGCACTTTGGGAGGTGACCTGTGTGAGTGCAGAGCCCTCAAG 2220
|||||
Db 2161 agaagttggcgaggcactttgggaggtgacctgtgtgagtgcaagacctcaag 2220
QY 2221 ATCTGTCAAGTCTTATTTCCCTTAATGTGTGATGAGGACACCTTTGCTGTGCTG 2280
|||||
Db 2221 atctgtcagcttcttgaatcccttaagtgtcgtlaagggaaaccccttgcgtgtagctg 2280
QY 2281 TCAAGATCTTACGCGCAGATGCCACCAAGAAATGCCAGTTCCTGCTCTCAGAGATG 2340
|||||
Db 2281 tcaagatctttagcgccagatgccaacaaatgccaagcttccctgttctccagaaatg 2340
QY 2341 ATTTCTGAAGAAGGTGAAGATCATGTGAGGCTCAAGAGCCCAACATCTTGGCTGC 2400
|||||
Db 2341 atttctgaagaaggtgaagatcatgtcgaggtctcaagaccccaacatcatctgcgtgc 2400
QY 2401 TGGGCTGTGTGCAAGAGACGACCCCTCTGATGATTACTGACTATGAGAAAGCGG 2460
|||||

XX Claim 1; Page 26-30; S1Pp; English.
 CC CDNA derived from tumor metastatic tissue was amplified using
 CC primers (given in AA084783-84) based on sequences (AA071101, AA071103)
 CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
 CC identified in an isolated subclone. The 3' sequence of PTK22 was
 CC obtained by reverse transcription (using the primer of AA084786) and
 CC PCR amplification (primers AA084787-88) of RNA of human breast
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
 CC is given in AA084782.
 XX
 SO Sequence 3754 BP: 713 A; 1145 C; 1121 G; 775 T; 0 other;

Query Match 91.3%; Score 3617.6; DB 16; Length 3754;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 3728; Conservative 0; Mismatches 19; Indels 29; Gaps 7;

QY 181 CCGGGGTCGAGCGCTGGCTGCGGAGAGAGCGATGAGAGGTGTGAAAGTGGCAT 240
 Db - 2 ccgggtcgagccgctggtgtctccggaaagagcgtatgaggtgtctcgaaggtgcat 61
 QY 241 TCACCTGAGCGATGGGTTGACTTGAAGAAATGCCAAGAGATGCTGCCACCCTCTTA 300
 Db 62 tctactgacgatagggtgtgacttgaagaaatgccaagatgtctgcccaacccctta 121
 QY 301 GGCCCGAGGATCAGAGCATGTGGAGACCAAGGCCCTCATCTTACTGCTGCTCT 360
 Db 122 gggcccgagagatcagagacatgagaccagagccctgtcatcttactgtctgctct 181
 QY 361 TGGTGGCAAGTGAAGATGCTGACATGAAGGACATTTGATCTGCTGCAAGGCCCTTNG 420
 Db 182 tgggtgcaagtgagagatgctgacatgaaggaacatttgcctccgaaggtccgcatg 241
 QY 421 CCTCGGCGATGACAGCGGACCATCCAGACAGTGAATCTGCTTCAAGCTCTGCT 480
 Db 242 cctcggcgatgacagacgagaccatccagacagtgacatctcgtctccagctcctgt 301
 QY 481 CAGATTCACATGCGCGCGCCGACAGAGGTGGAGACAGTGAAGGAGTGGGCTGTGT 540
 Db 302 cagattccacgc 361
 QY 541 GCCCGGAGGATCGGTGTTTCCCAAGAGAGAGAGTACTTGCAGTGGATTAACAAC 600
 Db 362 gcccgagaggtcggtgttcccaagagagaggtacttgcaggtgtgataltaacacgac 421
 QY 601 TCCACCTGTTGGCTCTGTGTGGACACCCAGGACGGCATGCCGGGGGCTGGCAAGAT 660
 Db 422 tccacctgtgtgtctgt 481
 QY 661 TCTCCCGAGCTACCGGCTCGTTACTCCCGGAGTGGTCCCGGTGATGGGCTGGAAAG 720
 Db 482 tctcccgagctaccggtctgttaccctccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 541
 QY 721 ACCGCTGGGTCAGAGAGTATCTCAGGCAATGAGACCTGAGGAGTGTGCTGAAG 780
 Db 542 accgctgggtcagagagtgatctcaggaatagagaccttgaggtgtgtgtgtgtgtgt 601
 QY 781 ACCTGGGCCCCCAGTGTGCCCCGACTGCTGCTTCAACCCCGGGGTCAGCGGTC 840
 Db 602 acctgtggcccccaagt 661
 QY 841 TGAGTGTCTGTGGGGTGAAGTGTGATGCTGCTGCTGAGGAGTGGAGTCTCTGCT 900
 Db 662 tggaggtctgtctgt 721
 QY 901 ACAACCCCTGTGTGGGACACAAATGATTTATCTGAGGCGCTGTACTAAAGACTCA 960
 Db 722 acacccctgt 781
 QY 961 CCAATACGAGACTACCGTGGGAGCTGCAATGAGGGGTCTGGGCCAGCTGGCAGATG 1020
 Db 1021 gttgt 1080
 Db 842 gttgt 901
 QY 1081 ACTATGTGGAGTGAAGCAACACACACTTCTCCAGTGGCTATGTGAGATGAGTTGAGT 1140
 Db 902 actatgtggagtgaagcaaacacacttctccagtggtatgtgagttgtgtgtgtgtgt 961
 QY 1141 TTGACCGGCTGAGGCGCTTCCAGGCTATGAGTGCATGCTATGACATGACACAGCTGTG 1200
 Db 962 ttgaccggctgagggcgcttccaggtatgctgaggtgttcaatacagacagctgtgt 1021
 QY 1201 GAGCCGCTGTGCTGCGGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 1022 gagccgctgt 1081
 QY 1261 AGGGGAGCCCATGCGCCCAACCTTAGGGGCACTGAGGGGACCCAGAGCCGGGCTG 1320
 Db 1082 aggggagcccatgcgcccaacctagggggaacctgtgggagcccgagccgggtgtgt 1141
 QY 1321 TCTCAGTCCCGCTTGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 Db 1142 tctcagtcctcctgt 1201
 QY 1381 GGCCCTGCTACTTTCAGCGAAATCTCTCATCTGTGATGTGATGAAATCTCTCTC 1440
 Db 1202 ggccctgttactcttcaagcgaatctcttcaatctctgaatgtgtgtgtgtgtgtgtgt 1261
 QY 1441 CGGCACTGGAGGACCTTCCCGCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 Db 1262 cggcaactggaggaaccttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1321
 QY 1501 ACTTCAGAGCTTGAAGTGGAGCCAGCCAGGCGCAGC---AGCCGCTGGCCAGGCGCAGG 1557
 Db 1322 acttcagagcttgaagtgagtgagccagagccagccagccagccagccagccagccagcc 1381
 QY 1558 GGAAGCCGACCGCATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617
 Db 1382 ggaagccgacgcacatctcttcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1441
 QY 1618 TCATTGCCCTATGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1677
 Db 1442 tcatlgtccctatgtctgt 1501
 QY 1678 GGGTGTGGAAGAGAGCTGACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1737
 Db 1502 ggggtgtggaagagagctgacgcttcatctctgtgtgtgtgtgtgtgtgtgtgtgtgt 1561
 QY 1738 ACAACCCCGCAGTCTGAGAGAGCCACCCCGTACAGAGAGCCCGGCTCTGTGGATTC 1797
 Db 1562 acaaccccgaggtctcagagagcaccaccccgtaacagggccccgcctcgtgtgtgtgt 1621
 QY 1798 CGCCCACTCCGCTCCGCTGT 1857
 Db 1622 cgcccaactccgc 1681
 QY 1858 ACCGCTCTTCTTGTGGCACTTACGCGCGCTCCCGTCAAGGCGGCGGCGGCGGCGGCG 1917
 Db 1682 accgctcttctgt 1741
 QY 1918 CTGTGGCAACCCACCAACACCCAGGCTCAAGTGGGAGTATGAGAGCTGAGAAAG 1977
 Db 1742 cctgtgtgcaaaccaacacacacacaggtctatgtgtgtgtgtgtgtgtgtgtgtgt 1801
 QY 1978 CAGGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2037
 Db 1802 caggcgcccgcttctgt 1861
 QY 2038 ACATTGTAACTGACAGGAGGTCAACGGGGGCAACATATGCTGCTGCTGCTGCTGCTG 2097
 Db 1862 acattgttaacctgt 1921

OY	2098	CAGGCGACGTCGGGGAATGGGCCCCCCCCAAGATGGATTTCCTCGTGAATCTGACACTCGGCTTCA	2157
Db	1922	caaggagcaagctcgaggagacgagggcccccaagatggaattccctctgatactcgaaccgcgttca	1981
OY	2158	AGGAGAAGCTGTGGCGAGAGGCGCACTTTGGGAGGTGCACCTGTGTAGAGTGCAGACGCTTC	2217
Db	1982	agggagaagctctggcagagggccaaatttggggagaggtagacctggtgagaggttcgacagccctc	2041
OY	2218	AAGATTCGTGTCAGCTTGAATTTCCCTTAAATGTGCGTAAGGAGACACCTTTGCTGGTAG	2277
Db	2042	aagatctggttagctctgattctcccttaatgtagtgtaagagacaccccttctgtgtag	2101
OY	2278	CTGGCAAGATCTTAAAGGCGCGAGTGGCCACGACGAAGAAGAGCCAGCTTCCTGTTCTCCACGA	2337
Db	2102	ctggcaagatccttaacgagccgaatgtpaaccaagaatg-----ccagaa	2143
OY	2338	ATGATTTCTGTGAAGAAGGTCTAAGATCATGTGCGGTCAGAGGCCCAACATCATTTCCGC	2397
Db	2144	atgattctccgaaagaggtggaagatacatgtcgaggtgcgaagaccacaacatcatctgcg	2203
OY	2398	TGCTGGGCGTGTGTGTGTCAGAGAGCAGACCCCTCTGCATGATTACTGACTAATGAGAAAG	2457
Db	2204	tgcctggcgctgtgtgtgcagagacgacccctctgatacttaactaatacvtgagaaag	2263
OY	2458	GGGACCTCAACCAAGTTCTCTAGTGGCCACACAGCTGGAGGACAAAGGACACCGAGGGGGGCC	2517
Db	2264	ggcgaccccaacacagcttctccatgagccacacagctggaagacaaggaacgcgagggggccc	2323
OY	2518	CTGGGAGCGGAGGAGGCTGGCGCAGAGGGGCCACACATCAGCTACCACTAGTGCTGCAATGTG	2577
Db	2324	ctggggagccgagcaaggtcgtgcgcaaggggcccaacatacgaatctgtctgcaatgtg	2383
OY	2578	CAGGCCAGATTCGCTTCGCGGATCGCTATTCGTGGCCACACTCAACTTTGTACATCGGAGCC	2637
Db	2384	cagccagacatcgctctccgagatgcatctgcgacacatacaactctgttaactcggagac	2443
OY	2638	TGGCGACGCGGAACTGCGCTAGTTGGGGAAATTTACCATCAAAATCGACAGACTTTGGCA	2697
Db	2444	tggcgacagcggaactgcctcagctgttggggaaaattcacaataatcgcagaacttctgca	2503
OY	2698	TGAGCGCGAACTCTATGCTGGGGACTATTACCGTGTGCGAGGGCGGGCAGTGGTCCCA	2757
Db	2504	tgaagccgaacaccttaatgtctgaggaactataacgtgtgagagggccggagctgtgccca	2563
OY	2758	TCCCGCTGGAATGGCCTGGGAGTGCATCTCTAATGGGGAAGTTCAAGACTGCGAGTGAACGT	2817
Db	2564	tcgcgtgaaatggccttgggaatgtgatactcaatgaggaatgtccacgcctgcgaagtgcagct	2623
OY	2818	GGGCGTTTGGTGTGAGACCCCTGTGGGAGGTCGTAGTGTGTGTAAGGCGCCAGGCCCTTGGGC	2877
Db	2624	gggccttctgtgtgaaacctgtgggaaggtgcataatgtcgtatgctgtatgggcccagcccttgggc	2683
OY	2878	AGCTCACCGACGAGCAAGGTATCGAGAACGCGGGGAGTTCTTCGGGACCAAGGCGCGGC	2937
Db	2684	agctcacacgagcagagcaggtatcatcgaagacaagcggggaagtcttccggagaccaagggccgc	2743
OY	2938	AGGTGTACCTGTCCCGGCGCGCTCGGCTGGCCCGCAGGGCCTTAATAGCTGAATGTTCCGT	2997
Db	2744	agggtatcctgtcccgccgcgcctgcctggccgagggccataatgagtctgactgtcgt	2803
OY	2998	GCTTGGACCCGGGAGTGTGAGACAGACGACACCCCTTTTCCAGCTGCATGGTTCCTGAGAG	3057
Db	2804	gcttggagccgggagctctgagcagcgacacccctttccagtgatcatgtgttctctggag	2863
OY	3058	AGATGACATCAACAGGCTGTGAATATACATCACTCAGACTGCCCTCTCCCTCAGAGGATGATC	3117
Db	2864	agatgtgactcaacaacaggtgtgataacacatcacagcttgcctctcccaaggagcgatc	2923
OY	3118	CAGGCGAAGCGACGATCACTAAACAAAGAGGACACAATAGGACGCTGTG-CCCTTCGCCCTC	3176
Db	2924	caaggagagcagctgaaactaataacaagagggacacaatgacacccctctgcctcttcccttc	2983

QY	3177	CCACAGGCCATCACACTCTTAATATAGAGGGCAGTGAAGTCTGCAGGTGGGCTGGGCGCCACCCAG	3236
QY	3178	CCACAGGCCATCACACTCTTAATATAGAGGGCAGTGAAGTCTGCAGGTGGGCTGGGCGCCACCCAG	3237
Db	2984	CCGAGAGCCCATACCTCACTCAATAGAGGAGTGAAGTCTGCAGGTGGGCTGGGCGCCACCCAG	3039
QY	3237	GGAGCTGATGACCCCTCTTCTCCCTCTTCTTGAGACACACTCTATGCTTCCCTCTCTGTTCTTC	3296
Db	3040	ggagctgagccctcttcctccttccttggaacaacatctcatgtcccttccttccttccttc	3099
QY	3297	TTCTTAGAAGCCCTGTGTGCCCAACCCAGCTGTGTCTGTGTGGATGCTCTTCACACCTTC	3356
Db	3100	ttccttagaagccctctgtcgcaccacagctgtgtctcgttgatgagatgtctctctccacccac	3159
QY	3357	CTCTAGGCATTCCTTTGGGGAGGAGGTGGGAGAGAAATTAGATATAGACATGACATGAGCC	3416
Db	3160	ccctctgacatccctcttggggaaggttggggaagaataataagataagacatctgacatgtgcc	3219
QY	3417	ATTGAGACACCTGGGGCCCACTGTGACATCATGTATCTCTGTGAGAGGTGGCTGGC-CCCA	3475
Db	3220	atttgagacaccttgggcccacatctgagaaacaatctgattccctggaaggttggctgtgccc	3279
QY	3476	GCTTCTCTTCCTGTACACACTGGACCCCACTGGCTGAGATCTGTGGGGTGAGAGGA	3535
Db	3280	gctctctctcctgttcaacacatctgaccccaactgtctgagatcttgggggttgagagga	3339
QY	3535	CAGAAGAGAGAGAAATTTCTCTTGTCTGTCTCTCTACTTGTCTCTCAGCTTGCGCT	3595
Db	3340	caagaagagagaaatgttctctgttctcgtctcttactgtctctcaagcttggct	3399
QY	3596	TCTTCTCTCATACCTGGAACACTGGACCTGGGGGTAGCCCGCCAGCCCTCAGT	3655
Db	3400	tcttctctccatcaacctgaaacaatctggagcttgggttgggttgggttgggttgggttgggtt	3459
QY	3656	CA-CCCCACCTTCCACTTGCAGTCTTGTAGTAGACTTCTCTAAGCCTATACGTTTCT	3714
Db	3460	cacccccaacttcccaacctgagctcttggcttgaacttctcttaagctataagttct	3519
QY	3715	GTGAGATGAATTTTGGGATTTGGGGGAGAAAGAGGAGCAACGGCCCTAGCCTTGGGGTTG	3774
Db	3520	gtgagatgaatatgttgatcttggggttggggaaggaaggaaggaaggaaggaaggaaggaagga	3579
QY	3775	GACATCTGTAGTGTAGCTCCACATTTGATTTTCTTATATCACTTGGGGTGTACATTT	3834
Db	3580	gacatctgtagtgtagcttgcacatgtatcttctataatcaact-gggtctgtacatct	3638
QY	3835	TTGGGGGGAGAGACACAGATTTTTCACATATATATGACATGCTTGAGGCAATTTTAA	3894
Db	3639	ttagggggagagagacagatctttaaacaataataatgagccctgagcttgggccaatttaa	3698
QY	3895	TCCCCTGTGACTGAGGAGGATTAATATTAAGGTTGTGATTTTCACAAAAA	3950
Db	3699	tcccctgtgactgagaggaataataaagtttgagtttccacaaaaa	3754
RESULT 5			
AA	16842	AA	16842
ID	AA	ID	AA
AC	AA	AC	AA
XX	AA	XX	AA
XX	14-FEB-2002	XX	(first entry)
DT	Human epithelial discoidin domain receptor 1 (YSG5) trke cdna.		
DE			
XX			
KW	YSG5: YSG5: schizophrenia; chronic animal model; LGSU: netrin receptor; LGSU:		

```

FT      /tag- a
FT      /product- "Human trke polypeptide"
XX      WO200175440-A2.
XX      11-OCT-2001.
XX      02-APR-2001; 2001WO-GB01486.
XX      31-MAR-2000; 2000GB-0007880.
XX      26-MAY-2000; 2000GB-0012768.
XX      (WELF-) WELFIDE CORP.
XX      Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
XX      WPI: 2002-010813/01.
XX      P-PSDB: MAU10542.
XX      Novel chronic animal model of schizophrenia, useful for identifying
XX      anti-psychotic drugs and genes that are associated with schizophrenia
XX      Claim 1; Fig 7a; 79pp; English.
XX      The invention relates to YSG polynucleotide fragments for use in
XX      diagnosing and/or developing treatments for schizophrenia using chronic
XX      animal models. The polynucleotides and their encoded polypeptides are
XX      used for identification of compounds which modulate the expression of YSG
XX      sequences, leading to the manufacture of schizophrenia medications. The
XX      sequences can also be used for testing candidate compounds for any effect
XX      on the polypeptides. Anti-schizophrenic effects of a compound can be
XX      determined by measuring local cerebral glucose utilisation (LCGU) or
XX      comparing its expression level with that of a control group. The
XX      sequences are useful in the identification of genes associated with
XX      schizophrenic states and in the development of an antibody. The sequences
XX      of the invention include phosphodiesterase 1-alpha, calcium-independent
XX      alpha-latrotoxin receptors (CIRL)-1.2a3, epithelial discoidin domain
XX      receptor 1 (trke), netrin receptor (UNC5H1), synapsins 1a and 1b and
XX      tumour necrosis factor (TNF) alpha. This sequence represents human
XX      epithelial discoidin domain receptor 1, trke (YSG5) DNA.
XX      Sequence 3554 BP; 682 A; 1064 C; 1065 G; 743 T; 0 other:
XX
Query Match      85.4%; Score 3383.2; DB 24; Length 3554;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 3545; Conservative 0; Mismatches 3; Indels 130; Gaps 3;

```

```

QY      637 ATCCGGGGGCTTGCGCAAGAGACTTCTCCGGAGCTACCGGCTGCTTACTCCGGAGT 696
DB      atgcgggggcttgccgcaaggagttctccggggtacacggcgctgtctctccggagtg 421
QY      697 GTGCGCGGTGAGATGGGCTGGAAGACCGTGGGCTAGAGGTGATCTCAGCAATGAG 756
DB      gtccgctggatggtgcttgaaaggaccgtggtgtaagagtgatctcaggcaatgag 481
QY      757 ACCCTAGGAGAGTGTGCTGAAGAGACTTGGGGCCCCGANTGGTGGCGAGCTGCTGCT 816
DB      accctaggagatggtgctgaaaggaccttgggcccccatggtgcccagctggttcgct 541
QY      817 TCTACCCCGGGCTGACCGGGTATGATGTCTGTCTGGGGTAGAGCTATAGCTGCTG 876
DB      tctacccccggctgaccgggtgcatagtgctgtcttggtggttagagctctatgctgc 601
QY      877 TCTGAGGAGATGAGTCTCTCTTACACCGCCCTGTGGGCAAGACATGTTATCTG 936
DB      tctggaggagatgagctctcttcttaccacgccccctggtggcagaacaatgattatctg 661
QY      937 AGGCGCTTACCTCAACGACTCCACCTATAGAGGACATACCGGGCGGAGCTGAGTAC 996
DB      aggcgcttacctcaacgactccacctacatgagagacatcgtggcggacgcagatcg 721
QY      997 GGGGTCTGGGCCAGCTGCGAGATGCTGTGGGGCTGGATGACTTTAGGAAGATCAG 1056
DB      ggggtctggccagctgagatgctgtgtgtgggctgtgattaggaagatcag 781
QY      1057 ACTCGGGGTCTGGCGGCTATGACTATGTGGATGAGGCAACACAGCTTCTCAGTG 1116
DB      actcgcgggtctggcgccgtatgactatgtggatgaggaacacacgcttctcagtg 841
QY      1117 GCTATGTGAGATGAGTTGAGTTGACCGGCTAGAGGCTCTTCAGGCTATGAGTCTC 1176
DB      gctatgtgagatgagttgagttgacccgctagagggctcttcagggctatgagagtgctc 901
QY      1177 ACTGTACAAATGCAACACGCTGGAGCCCGTCTGCTGGCGGGTGAATGTCTCTTC 1236
DB      actgtacaaatgcaacacgctggagcccgctgctggcggggtgaatgtctcttc 961
QY      1237 GGGCGGGCCCTGCAATGGCGCTGGAGGGAGGCCATCGCCACAACTAGGGGGCAAC 1296
DB      gggcgggccctgcaatggcgcctggagggagggccatcgccacaactagggggcaac 1021
QY      1297 TGGGGAGCCCGCAGAGCCCGGGCTGCTCAGTACGCCCTTGGGGCGGCTGTGGCTGCTTC 1356
DB      tggggagcccgagagcccgggctgctcagtacgcccttggggcggctgtggctgcttc 1081
QY      1357 TGCAGTCCCGCTTCTCTTTGGCGGGCCCTGATTACTTTCAGGAAATCTCTTCATCT 1416
DB      tgcagtcccgcttctcttttggcgggccctgattactcttcaggaatctcttcactct 1141
QY      1417 CTGATGTGATGAACAATTCCTCCGAGCACTGGAGAGGACCTTCCCGACACCCCTGTGT 1476
DB      ctgatgtgatgaacaattcctccgagcactggagagaccttcccgacacccctgtgt 1201
QY      1477 GGGCGCTTGGGCCACTTCCACCACTTCAGACAGCTTGGAGCTGAGGCCAGAGCCAGC 1536
DB      gggcgcttgggccacttccacacttcagacagcttggagctgagggccagagccagc 1261
QY      1537 AGCCCGTGGCAAGGCGAGGGAGAGCCGAGCCGCAATCTCATGCGCTGCTGTGGCA 1596
DB      agcccggtggcaaggcgagggagccgagccgcatctcatctcgtgctgtgtggca 1321
QY      1597 TCATCTGCTCTCTGCTCTCTTCAATGCTTCATAGCTCTTGCGGCTGCACTGGCGCAGGC 1656
DB      tcatctgctctctgctcttcaatgcttcatagctcttgcggtgcaactggcgagcg 1381
QY      1657 TCTCAGCAAGGCTGAGAGGAGGTGTGGAAGAGAGCTGACAGCTTCACTCTCTGCTGC 1716
DB      tctcagcaaggctgagagaggtgtggaagagagctgacagcttcaactctctgctgc 1441
QY      1717 CTGGGGAACACTATCTCATCAACACCGCCAGGTCTCTAGAGAGCCACCCCGTACAGG 1776

```

Db 1442 ctgggacacatctcatcaacaacagccaggtccctagagagccaccccccagcag 1501
QY 1777 AGCCCGGCGCCGCGGGAAATCCGGCCACATCCGCTGCTGTCGCCCAATGGCTCTCGT 1836
Db 1502 agcccgcgcgcgctggaaatccgcgccacatccgctccctgctgctcccaatggcct--- 1557
QY 1837 TGCTGCTCTCCAAATCCAGCCTACCGCCTCTGCTGCGCACTTAAGCCCGTCCCTCGAG 1896
Db 1558 ----- 1557
QY 1897 GCCCGGGCCCCCACCACCCGCTGGGCAAAACCACACCCAGGCTTACAGTGGG 1956
Db 1558 -----gctcaagctggg 1570
QY 1957 ACTATATGAGACCTGAGAAAGCCAGCCGCCCGCTTCTGCCCCACCTCCCAAGACGCG 2016
Db 1571 actatactgagacctgaaagacagcgcccgctctctgccccacatccccaagacagcg 1630
QY 2017 TCCCCCAATTATGCGAGGCTGACATTTGTTACCTCGACAGGGGTACCCGGGGGCAACACT 2076
Db 1631 tcccccttatgcccgaagctggaactgttaacctgcaaggcgctcaaccggygcaacacct 1690
QY 2077 ATGCTGTGCTGACACTGCCCCCAGGGGCACTGGGGATGGGCCCCCAAGATGATTTCC 2136
Db 1691 atgctgtgctgcactgcccccaaggcgagctcggygaltggygcccccaagatgttcc 1750
QY 2137 CTGCATCTCAGATCCCGCTTCAAGAGAAAGCTTGGCGAGGGCCANTTGGGGAGGTGCACC 2196
Db 1751 ctgcattctgcactcgcgtctcaaggaaagcttgcgagggccagtttgggaggtgaccc 1810
QY 2197 TGTGTGAGGTGACAGCCCTCAAGATGTGTGATCTTGAATTTCCCTTAATGTGCTA 2256
Db 1811 tgtgtgaggttgacagccctcaagatctgttcagctcttgatcttcccttaatgtgtgta 1870
QY 2257 AGGAGACACCTTTTCTGCTGTAAGCTTAAAGATCTTACGGCAGATGCGACCAAGATGCCA 2316
Db 1871 agggacacaccttctgtgtgactgttcaagatcttaacgagcagatgcacacaaagatg--- 1927
QY 2317 GCTTCTCTGTTCTCTCGAGGAATGTTCTCGAAGAGGAGGAAATGATGCGAGGCTCA 2376
Db 1928 -----caggaatgalttccgaaagagtgaaatcaatgctcgaggtca 1972
QY 2377 AGGACCCCAACATATTCGGCTGCTGGGCGTGTGTGTGACAGAGACCCCTCTGCATGA 2436
Db 1973 aggacccaacatcttcgctgt 2032
QY 2437 TTACTGACTACATGAGAGAGCGGCACTCAACAGTTCTCTCACTGCTCCACACAGCTGAGG 2496
Db 2033 ttactgactacatgagagacgagcactcaaacagttcttcaagctgccccacagctgag 2092
QY 2497 ACAAGGAGAGCGAGGGGGCCCTGCGGAGACGGGCGGCTGGGCGGCGGCGGCGGCGGCT 2556
Db 2093 acaaggaagcgcgaagggccccctgtggagcgaggtcgctgcaagggtcccaacatcgct 2152
QY 2557 ACCCAATGCTGCTGATGTGTGACAGCCAGATCGGCTCGGGGATCGCTATCTGGCGCAC 2616
Db 2153 acccaatgctgtgcatgt 2212
QY 2617 TCAACTTTGTAACATCGGAGACCTGGCCACGCGGAATGCTGCTGTTGGGAAAATTTACCA 2676
Db 2213 tcaactttgtacatcgagactgt 2272
QY 2677 TCAAAATCGAGACTTGGGCTGAGCGGAACCTCTATGCTGGGGATATTATACGCTGTC 2736
Db 2273 tcaaaatcgagacttggcatgt 2332
QY 2737 AGGCGCGGCGAGTGTGCTGCCATCCGCTGATGCGCTGGAGTGCATCTCATGGGGAGT 2796
Db 2333 agggcgcggagctgt 2392
QY 2797 TCAGAGATCGAGATGAGCTGTGGGCTTTGGTGTGACCTGTGGGAGGTGCTATGCTCT 2856
Db 2393 tcaagactcgagatgagctgt 2452
QY 2857 GTAGGGCCCAAGCCCTTTGGGACAGCTCACGACGACGAGTCAATCAGAGCGGGGAGT 2916
Db 2453 gttagggcccaagcccttgggacagctcaacgagcaggttcaatcgaaagcgaggtgtgtgt 2512
QY 2917 TCTTCCGGGACCAAGGCGGCGGACAGGTGTAACCTGTGCCCGGCGGCTGCTGCCCGAGGGCC 2976
Db 2513 tcttcggyagacagggcgccaggtgtactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2572
QY 2977 TATATGAGCTGATGCTTGTGCTGTGAGCCGGGAGTGTGAGACGACACACCTTTTCC 3036
Db 2573 tataatgactgt 2632
QY 3037 AGCTGCATCGGCTTCTGCGACAGAGATGACTCAACAGGTGTGATATACATCCACTG 3096
Db 2633 agctgcactgcttctctgtgcaagagatgacatcaacacggtgtgtgtgtgtgtgtgtgtgt 2692
QY 3097 CCCCCTCCAGGAGATGATCCAGGGAGCGACATGACACTAAATAAAGAGAGACAAATG 3156
Db 2693 cccccctcaagggaggtgtatccaggggaagcagtgacatcaaaacaaagagacaaatg 2752
QY 3157 GCACCTTGCCCTTCCCTCCCGACAGCCCATCACTCTAATAGAGGACAGTGAATGCA 3216
Db 2753 gcaccttgcccttccctcccgacagcccatcaaccttaatagagcagtgagactgca 2812
QY 3217 GGTGGGCTGGGCCACCCAGGAGAGCTATGCCCTTCTCCCTTCCCTGACACACTCTCA 3276
Db 2813 ggtgggctggggccacccagagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2872
QY 3277 TGTCCCTTCTGTTCTTCTCTCTTAAAGCCCTGTGCGCCACCGACGAGTGTGCTGTG 3336
Db 2873 tgtcccttctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2932
QY 3337 ATGGAGTCTCTCCACCTCTCTAGCCATCCCTGGGGAGAGGTGGGAGAAATATAGG 3396
Db 2933 atggagtccttccacccct 2992
QY 3397 ATAGACATGAGACATGAGCCATTGAGACACCTGGGCGCCACTGGAACACTGATCTCTG 3456
Db 2993 atagacactgagacatgagccatgtgagacacccgtggcccaactgtgacaacactgtctct 3052
QY 3457 GAGAGTGGGTGCG-CCCGAGCTTCTCTCTCCCTGTACACACTGAGACCCACTGCTGA 3515
Db 3053 gagagtggt 3112
QY 3516 GAATCTGGGGTGAGAGACAAAGAGAGAGAGAAATGTTTCTTGTGCTGCTCTGT 3575
Db 3113 gaatctgggggtgagagacaaagagagagaaatgttctctgtgtgtgtgtgtgtgtgtgt 3172
QY 3576 ACTTGTCTGAGCTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3635
Db 3173 acttgtctgagcttgggcttct 3222
QY 3636 GCCCGCGCCAGCCCTAGTCACCCCACTTCCCACTTGGAGTCTTGTAGTACACTTC 3695
Db 3233 gccccgcgcagccctcaagtcacaccccaatcccaactcagctgtgtgtgtgtgtgtgtgt 3292
QY 3696 TCTAAGCTTATGCTTCTGT 3755
Db 3293 tctaaagcttaactgttctgt 3352
QY 3756 GCCCATGACCTTGGGGTGTGACATCTAGTGTGCTGCGCACATTTGATTTTCTATATC 3815
Db 3353 gcccatatgcttgggt 3412
QY 3816 ACTTGGGTTTGTACATTTTGGGGGAGAGACACAGATTTTAACTAATATATGAGAC 3875
Db 3413 acttggggttgtacatcttctgt 3472
QY 3876 TAGCTTAGGCAATTTAAATCCCTGTGACTAGAGAGATATATTAAGGTTGATTTCC 3935
Db 3473 tagcttaggcaatttaactccctgtgactagagcaagtaataaaggtgtgtgtgtgtgtgt 3532

Db	1021	atgggccccccagggatggtattccctccgatctccgatccgcgtctcaaggagaagcttgcg	1080
OY	2173	AGGGCAGATTGGGGAGGTGCACCTGTGTAGGTGACAGCCCTCAAGATCTGCTCACTC	2232
Db	1081	aggggccagtttgggggaagtccactctgtgaggtgcacagccctcaaatcctgtgcagctc	1140
OY	2233	TTGGATTTCCCCCTTAATGTGCGTAAGGAGACACCCCTTGGTGTACTGTGCAAGATCTTAC	2292
Db	1141	ttgatctccccccttaatggtgcgaaggagaaacctctgtctgacgtctcaagaacttaac	1200
/OY	2293	GGCCAGATGCCACCAAGAAATGCCAGGCTTCTCCTGTTCACGGAATGATTTCTGTGAAG	2352
Db	1201	ggccagatgcccacaaagaatg-----ccaggaatgattccctgaag	1242
OY	2353	AGGTGAAGATCATGTGCAAGCTCAAGACCACCAATCATTTGGCTGCTGGGCGTGTGTG	2412
Db	1243	agggtgaagatcatgctcgaaggtctcaaggaccacaaactcaatctggtctgggcgtgtatg	1302
OY	2413	TGCAGAGAGACCCCTCTGCTGATGATTACTGATCTACATGTGAGAAACGGCAGACTCAACAGT	2472
Db	1303	tgcagagagacccccctctctgattgattactgactcaatctggagaaacggtgactcaacagct	1362
OY	2473	TTCTTCATGCCCCACACAGCTGTGAAGAACAAAGCGACGAGGGGGCCCTTGGGACGGGCAAG	2532
Db	1363	ttccctcatgccccccacccagctctggagaaagacagccgagggggcccccttggggacgggcag	1422
OY	2533	CTGGCAGAGGGGCCACCATATAGCTACCCCAATCTGCTGATCTGCGACCCAGATCGGCT	2592
Db	1423	ctggcagagggccccacatcagctaccatgctcgtgcatgtgtgcagcccaagatgcct	1482
OY	2593	CCGGCATGCGCTATCTGCGCCACCTCACTTTGTATCAGTGGGACCTGGCCACGCGGAAT	2652
Db	1483	ccggcatgcgctatcttgcgcacactaaacttgtacatcgtgacctggccacgcggaact	1542
OY	2653	GCTAGATGGGGGAAAATTTACCATATAAATGCGAGACTTTGGCATGAGGCGGGAACGCT	2712
Db	1543	gctcagatctggggaaaatttcaacatcaaaaatcgcagactcttgccttgagccggaaacctct	1602
OY	2713	ATGCTGGGGACTATTACCGTGTGACAGGGCGGGGAGTGTGCGCATCCGCTGATAGGCGCT	2772
Db	1603	atgctggggactattccgcgtgcagggccgggcaggtgcctgcacatccgtctgctgcct	1662
OY	2773	GGGAGTGCATATCTCATTANGGGGAATTTACACATCGGACAGTGAAGTGTGGGCTTTGGTGTGA	2832
Db	1663	gggagtgcatccctccatctggggaaatttcaacgactcgcagatgtaacgtctggtgaccttggta	1722
OY	2833	CCCTGTGGGAGGTGCTGATGCTCTGTAGAGGCCACACCCCTTGGGAGCGTCAACACAGC	2892
Db	1723	ccctgtggaggtgctcgtatgctctgtgaaaggccccagcccttgggagatcaccgaaagac	1782
OY	2893	AGGTTCATCGAAGACGGGGGGGAGTTTTCGGGAGACAGGGCGGCGAGGTGTAACCTGTGCC	2952
Db	1783	aggttcatacgaagacggggggagtttcttcctggagacaaaggccgggaagttgtaacctgtccc	1842
OY	2953	GGCCGCTGCTGCTCCCGGAGGGCGCTTAATGACTATATCTTCGTGGTGTGGAGCGGGAGT	3012
Db	1843	ggccgctgctgctcccgca-ggcytatatgaaactgtatgctcttcgtgtcctggagccggaggt	1901
OY	3013	CTGAGCAGCGACACCCCTTTTCCCACTGCATCGGTTCTCTGGACAGAGTACTCAACA	3072
Db	1902	ctgagcagcgacacacctttctccagctgtcaatcgttctctgcgcagagatgtaactcaaca	1961
OY	3073	CGGAGTAAATCAACATCCAGCGTGGCCCTCTCAAGGAGTGAATCCAGGGGAACCAAGT	3132
Db	1962	cggatgtatcatcaatcatcaccgcgtgcctccctccctccagggagatcccaagggaagccaggt	2021
OY	3133	ACACTAAACAGAGAGACAAATGSCACTGTGCCCTTCCCTCCGACACCCATCAACC	3192
Db	2022	acactaaaaacagagacacaatgtgaactctgtcccttccctcccgacagccatacac	2081
OY	3193	TCCTAATGAGAGGTAGACTGCAAGTGGGCTGGGGCCACCCAGGAGCTGATGCCCTT	3252

[illegible]

PR 03-NOV-1999: 9905-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI: 2001-235357/24.
DR P-PSDB: AAG73767.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
1
XX Claim 1: Page 2388-2389; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Additionally, N may be used to produce the patient's own production of P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 2861 BP; 567 A; 899 C; 800 G; 589 T; 6 other;

Query Match 70.3%; Score 2783.6; DB 22; Length 2861;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2833; Conservative 5; Mismatches 4; Indels 20; Gaps 3;

QY 1093 GGAGCAACCAAGCTTCTCCAGTGTGAGATGAGAGTTGAGTTGACGGCTGA 1152
Db 1 ggcacacacagcttccacgtgctatgtagagatgtagttagttgacgcgtga 60
QY 1153 GGGCTTCCAGGCTATGAGAGTCACTGTAAACAATGACACAGCTGGAGCCCGCTGC 1212
Db 61 gggccttcacagctatgtagtccactgttaacaacatgacacagctggagccgctgc 120
QY 1213 CTGGCGGGGTGGAATGTGCTTCCGGCGTGGCCATGCGCGGGAGGGAGGCCA 1272
Db 121 ctggcggggtggaatgtgcttccggcggtggccatgcgcgggagggagggccca 180
QY 1273 TGGCGCAACCTAGGGGGCAACCTGGGGGAGCCCGAGCCCGGCTGTCTAGTGC 1332
Db 181 tggcgcaaacctagggggcaacctgggggagcccgagcccggtgtcttagtgc 240
QY 1333 TTGGCGGGCGGTGCTGCTTTCGAGTGCCTTCTCTTTCGGGGGCGCTGATTAC 1392
Db 241 ttggcgggcggtgctgcttcttcgagtgccgtctcttcttcggggcgctgttac 300
QY 1393 TCTTCAGGAATCTCCTTCATCTCTGATGTGTGTAACAATTCCTCCGGACATGGAG 1452
Db 301 tcttcaggaatctccttcacatctctgctgtagtgaacaatctccttcggacatggag 360
QY 1453 GCACCTTCCGCCAGCCCTGTGTGGCGCTGGGCCACTCCACCACTTACAGACACT 1512
Db 361 gcaaccttcgccagccctgtgtggcgctggggccactcccaacttcagacact 420
QY 1513 TGGAGCTGAGAGCCAGAGGAGGAGCCCGGTGGCCAGGCGGAGGAGCCGAGCCCA 1572
Db 421 tggagctgagagccagagagcgagccgtgtgcagagcgagagggagccgagccgaca 480

QY 1573 TCTTCATCGGCTGCTGTGTGGGCCATCATCTCTCTCTGCTGCTCATCATTTGCCCTATGC 1632
Db 481 tcttcacgtgctgctgtgtgggccatcatctctctctgctgctcatcatgtccatgc 540
QY 1633 TCTGGCGGCTGCACTGGCGAGGCTCCTCAGCAAGGCTGAAGGAGGAGGTTGGAAGAG 1692
Db 541 tctggcggtgctgcactggcgaggtcctcagcaaggctgaaggagaggtgtggaagag 600
QY 1693 AGCTGAGGTTCACTCTCTGTCTGCTGGGAGACATACTCATCAACAACCGCCAGAGTC 1752
Db 601 agctgaggttcaactctctctgctgctgggagacatactcatcaacaacccgagagtc 660
QY 1753 CTAGAGAGCCACCCCGGTACACAGAGCCCGGCTGTGGGAATCCGCCCATCCGCTC 1812
Db 661 ctgagagagccaccccggtacacagagcccgctgtggaatccgcccacatcgtc 720
QY 1813 CTGTGTCCCAATGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1872
Db 721 cctgtgtcccaatgagctgtgctgtgctgtgctgtgctgtgctgtgctgtgctgtg 780
QY 1873 CCACTTACGCGCTCCCTCGAGGCCCGGCCCCACACCCCGCTGGGCAACCA 1932
Db 781 ccaactacgctccctcgagcccgccgcccccaacacccgctggccaaccca 840
QY 1933 CCAACACCAAGGCTTCACTGTGGGAGTATATGAGCTGTGAAGCAAGCGCCGCTTC 1992
Db 841 ccaacaccaaggttcaactgtggaactatagagctgtgagagcagagccgcttc 900
QY 1993 TGGCCCGACCTCCCGAGAACAGGCTCCCATATGACAGGCTGACATTTGTTACCTGC 2052
Db 901 tggcccgacctcccgagaaagagctcccatatgacagagctgtgacatgtgtaac 960
QY 2053 AGGCGCTCAACGGGGGCAACACCTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2112
Db 961 aggcgctcaacgggggcaaacacctatgctgtgctgctgctgctgctgctgctgctg 1020
QY 2113 ATGGGCCCCCGAAGTGTGATTTCCCTCGATCTCGACTCCGCTTCAAGAGAACTTGGCG 2172
Db 1021 atgggcccccgagtggtatctccctcgatctcgactccgcttcaagagaaacttggcg 1080
QY 2173 AAGGCGAGTTGGGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2232
Db 1081 aaggcgagttggggagggagtggtgagtggtgagtggtgagtggtgagtggtgagtg 1140
QY 2233 TTGATTTCCCTTATGTGCTGAAGGAGACACCTTGTGCTGCTGCTGCTGCTGCTGCT 2292
Db 1141 ttgatttcccttattgtgctgaaggagacaccttgtgctgctgctgctgctgctgctg 1200
QY 2293 GGCAGATGCCCAAGAAATGCCAGCTTCTCTGTTCTCAGGAATGATTTCTGAAG 2352
Db 1201 ggcagatgcccagaagatg-----ccagaaatgatttccgaaag 1242
QY 2353 AGGTGAAGTCAATGTGAGAGGTCAAGAGCCCAATATGATTTGGTGTGGGCTGTGTG 2412
Db 1243 aggtgaagtcaatgtgagaggtcaagagcccaatattggtgtgtgtgtgtgtgtgtgt 1302
QY 2413 TGCAGAGCACCCCTCTGTCATGATTTACTGATGATGATGATGATGATGATGATGATG 2472
Db 1303 tgcagagcacccctctgcatgattactgatactatagtagagagcgagctccaacagt 1362
QY 2473 TCTCATGTGCTTCAAGCTGAGAGCAAGAGCAGCCGAGGGGCGCTGTGGGAGCGGAGG 2532
Db 1363 tctcatgtgcttcaagctgagagcagcgagggggcgctgtgggagcgagcgag 1422
QY 2533 CTGGCAGAGGGGCGCACACACAGCTACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTG 2592
Db 1423 ctggcagaggggcgcacacagctaccaagctctgctgctgctgctgctgctgctgctg 1482
QY 2593 CCGGATCGCTATCTGCGCACTCACTTGTGATGATGATGATGATGATGATGATGATGAT 2652
Db 1483 ccggatcgctatctgcgcaactcaacttgtatcatcgggactctgctgctgctgctgct 1542
QY 2653 GCCTAGTTGGGGAATTTCAACATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCT 2712

```

Db 1543 gctctgttggggaatttcacacacaaatgcagacttggcttgagccggaacctt 1602
Qy 2713 ATGGGGGACTATTACCGTGTGAGGGCGGGGAGTGTGCGCCATCGCGTGGAGGGCT 2772
Db 1603 atgcctggggaactataacacgttgcagggccgggagtcgcccacccgcggagggctt 1662
Qy 2773 GGGAGTGCATCCTCATGGGGAGATTACAGACTGCGAGTGAAGTGGGCTTTGGTGA 2832
Db 1663 gggagtgatcctcatctggggaagttcacgactcgagtgagtggtggccttggctgga 1722
Qy 2833 CCTGTGGAGAGTGTGATGCTCTGTAGGGCCAGCCCTTTGGGAGCTCACCGAGAGC 2892
Db 1723 ccttgggaggtgtgtgactgtctgtagggccagcccttgggaggtccacggcgggc 1782
Qy 2893 AGGTATGAGAGAGCGGGGGAGTCTCCGGGAGCAGGGCGGAGGTGATACCTGTCC 2952
Db 1783 aggtatcagaaacgggggaggtcttcgggagccagggccggaggtgataccggtccc 1842
Qy 2953 GGGCCCTGCTGCGCCGAGGGGCTATATAGCTGATGCTTGGTGTGAGCGGAGT 3012
Db 1843 ggcgcctgtcctgcccga-ggcycatatagtctgactgtcctgggtgcggcgggaggt 1901
Qy 3013 CTGAGCAGCAGACCCCTTTCCAGCTGCATCGGTCTCTGGCAGAGATGCATCAACA 3072
Db 1902 ctgagcagcagacaccccttccagctgcacgtctcctggcagagagatcacaca 1961
Qy 3073 CGGTGTGATATACATCCAGTGTGCCCCCTCCCTGAGGGAGTATCAAGGGAGCACTG 3132
Db 1962 cgtgtgatacaacataccagctgcctcctcagggagagatacagggagagcagtg 2021
Qy 3133 ACATAAAGAGAGAGACATAGGACACTGTGCCCTTCCCTCCGAGAGCCCATCAC 3192
Db 2022 aacataaagagagagacaatagcactctgccccttccctcccgagcccatcac 2081
Qy 3193 TCTATATAGGACAGTGAACACTGCAAGTGGGCTGGGCCACCGAGGAGCTATGCCCC 3252
Db 2082 tctaataagagcagtgagcagctcaggtggtggccacacagagagctgacccctt 2141
Qy 3253 CTCCCTTCCTGGACACACTGTCAATGTCCCTTCTCTTCCCTCTAAGAGCCCTG 3312
Db 2142 ctcccttccctggagacacactcactatgtccctcctgtcttccctcctaaagccctg 2201
Qy 3313 TCGCCACCCAGCTGTGCTGTGATGGGATCTCTCCACCCCTCTAGCCATCCCTG 3372
Db 2202 tcgcccacccagctgtgtcctgtgagatgcctcctccacccctcctcagcatccctgt 2261
Qy 3373 GGGAGGAGTGGGAGAAATATAGATAGACACTGACACTGGCCATTTGAGACCTGGCC 3432
Db 2262 gggagaggtggggaatataagataagacactgacatggccatlgagcacctgggc 2321
Qy 3433 CCCACTGACACACTGTATCTCTGAGAGTGGTGGGCG-CCCGAGCTCTCTCCCTGT 3491
Db 2322 cccactgtgacacactgtctcctggagaggtgtgtgcgcccacagcttctcctccctgt 2381
Qy 3492 CACACACTGACCCCACTGCTGAGAAATCTGGGGGTGAGAGGACAGAGAGAGGAAA 3551
Db 2382 cacacactgagacccacactgtgtgaaatctgggggtggagaggaagaagagagaa 2441
Qy 3552 ATGTTTCTTGTGCTGTGCTGTGATGCTGTGAGCTTGAGGCTTTCTCTCTCATCA 3611
Db 2442 atgttctctgtcctgtcctgtactgttctcagcttggcttcttccctccatca 2501
Qy 3612 CCTGAAACACTGAGACTGGGGGTAGCCCGCCGAGCCCTAGTACACCCCACTTCCAC 3671
Db 2502 cctgaaacactgagactgggggttagcccgccagccctcctcagaccccaactccac 2561
Qy 3672 TTGAGCTTGTAGCTAGAACTTCTTAAGCCTATAGCTTTCTGTGAGTAAATATTGG 3731
Db 2562 ttgagcttctgtagctagaactctcctaagcctatagcttctgtgagtaaatattgg 2621
Qy 3732 ATTGGGGGAAAGAGGAGCAAGCGCCATAGCTTGGGGTGGACATCTCTAGTGAAC 3791

```

```

Db 2622 attgggggaagagggagcaagcccatagccttgggtgttgacatcctagtgagc 2681
Qy 3792 TGCCACATTAATTTTCTAATAATCACTGGGTTTGTCATTTTGGGGGAGAGACACA 3851
Db 2682 tggccacttgatcttccataatcacttgggtgtgtacatcttgggggagagacaca 2741
Qy 3852 GATTTTACACTAATATATATGAGACCTAGCTTAGGCAATTTAATCCCTGCACATGCGAG 3911
Db 2742 gatcttcactaatatattggaacctagcttggagcaatttaattcctcctgacatggag 2801
Qy 3912 GTAATATTAAGGTGAGTGTTCACACAAAAA 3953
Db 2802 gtaataataaggttgagtttccacaaaaa 2843

```

RESULT 8

```

AAV48292
ID AAV48292 standard; cDNA; 3096 BP.
XX
AC AAV48292;
XX
DE 16-NOV-1998 (first entry)
XX
DE Discoidin domain receptor 2 gene.
XX
KW Discoidin domain receptor; transformation; metastasis; collagen; ss;
KW Cleidocranial dysplasia; Stickler syndrome; extracellular matrix; MMP-1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 354..2921
FT /tag= a
FT /product= "Discoidin domain receptor"
FT sig_peptide 354..416
FT /tag= b
FT mat_peptide 417..2918
FT /tag= c
XX
PN MO9834954-A2.
XX
PD 13-AUG-1998.
XX
PF 05-FEB-1998; 98WO-CA00093.
XX
PR 06-FEB-1997; 97US-0041578.
XX
PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
PI Pawson A, Vogel W;
XX
DR WPI: 1998-447168/38.
DR P-PSDB; AAM77114.
XX
PT Novel ligands of discoidin domain receptor tyrosine kinase,
PT especially collagen - useful for treating e.g. metastasis,
PT cleidocranial dysplasia or Stickler syndrome
XX
PS Disclosure; Fig 22a; 115pp; English.
XX
XX

```

The DDR can be used to identify and evaluate substances which affect DDR receptor tyrosine kinase signalling pathways in the cell. Compounds CC which modulate such signalling pathways can be used to alter CC transformation or metastasis in mammals, to treat conditions involving CC structural or functional deregulation of collagens, e.g. Cleidocranial CC dysplasia or Stickler syndrome, conditions requiring modulation of CC extracellular matrix synthesis, degradation or remodelling, or to treat CC conditions needing modulation of MMP-1 expression such as wound healing.

Sequence 3096 BP; 762 A; 791 C; 752 G; 791 T; 0 other;

Query Match 16.2%; Score 642; DB 19; Length 3096;

QY	1000	GTCTGGGCGCAGGCTGGCAGATAGTGTGGTGGGGCGCTGGAGTAACTTTAGAAAGTCAAGAGC	1059
Db	1054	--ctaggcccaattgaaccgaatggtgtgtctgtgcctgcgcagcatlcttaaccagaccatgaat	1111
QY	1060	TGCGGGTCTGGCCAGCGCTATGATGACTATGTGGATGGAGTGAACACACAGGCTTCTCCAGTGGCT	1119
Db	1112	aaacaagctgtgcccgcgtactatgactcatgttggctgtgcggaacgaagatgtcccaatgtgc	1171
QY	1120	ATGTGAGATGGAAGTTTGAAGTTTGACCGGCGCTGAGAGGGCGCTTCCAGGCTAAGAGGTCACCT	1179
Db	1172	acattggtacatgttctgaatttgaccgcatcagaaatttaactacatcattgaagttcaact	1231
QY	1180	GTAACAAATGACACACCGCTGGAGACCGCTGTGCTGGCGGGGTGGAAATGTCTTCGGCGC	1239
Db	1232	gcaacaacaatgtttgtctaaagggtgtgaagatctcttaaggagtgtaacatgtctacttcgcgt	1291
QY	1240	GTCGGCCCTGTGCATGAGCCTGTGGAGGGGAGGCGCCATGAGGCCACACCTTGAAGGGGCAACCTGG	1299
Db	1292	ctg---aagccagttagtgggaaaccttaatgcattctctccctctgtcctgtatgaag	1348
QY	1300	GGAGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTGGCGGCGGTCGTGCTTCCTTGTGC	1359
Db	1349	tcaaccccagtgctcggtttgttcaaggtgtgctctccaacacgaaatgycagtgcatca	1408
QY	1360	AGTGGCGGCTTCCTCTTTGGCGGGGCGCCGTGTTACTCTTACAGGAATTCCTTACTCTGTG	1419
Db	1409	agtgtaaatcccaatttgcacgaataccctggaatgtatgttcagttagatcaactcccaatcag	1468
QY	1420	ATGTGTGAGCAATTCCTCTCCGGCACTGGAGGACCTTCCGGCCAGCCGCCCTGGTGGC	1479
Db	1469	atgtctgaatgttcaacaacatctcgaagcctgtcccaactctcca-----	1513
QY	1480	CGCGTGGCCACCTCCACCCACTTGACAGACTTGAAGCTGAGAGCCAGAGGCGCAGAGC	1539
Db	1514	-----tgaaaccaacaactatgac	1534
QY	1540	CCGTGGCGCAAGGCGGAGGAGGAGCCGACCGCATTCCTCATACGGCTGCTGTGGCCATCA	1599
Db	1535	caatgtctaaagtcttgaatgaagaacaactcgaactcgaatgtgctgtgtgtgtgcatca	1594
QY	1600	TCTGTCTCGTGCATCATATTGACCCCTCATGTGTCGTGGGGGCGTGCACCTGGCGCAGAGCTCC	1659
Db	1595	tcttatctctctctgtgcacatcatgtcatcctctctgtgaagacgtctctgtgcagaataatgc	1654
QY	1660	TGAGCAGAGCTGAACGAGGAGGTGTGGAGAGAGAGCTGACAGGTTACCTCTGTGCCCTG	1719
Db	1655	tggagaagagctctctcggagagtgatgtgaatgaatgaagtcagacgtcttccctgcaca	1714
QY	1720	GGGACACTTCTCTCATCAACACCGGCCACAGGCTCTAGAGAGCCACCCCGCTACGAGAGC	1779
Db	1715	gtgatctctagcatgttcaacaat-----aacgcgtctct	1747
QY	1780	CCGCGGCTGTGGGAATTCGGCCCATCTCCGCTGCCGTGTGCCCAATGGCTCTGCGTTGC	1839
Db	1748	catcaactatgtgaacaaggtccaaactcgaactaagatcatgtcatctt-----	1794
QY	1840	TGCTTCCAATCCAGCTACCGGCTCTTCTGTGGCACTTACGCGCCCTGCCCTCGAGAGCC	1899
Db	1795	-----ccctctgcgc	1804
QY	1900	CGGGCCCCCACAACCCGCTGAGGCCAAACCAACCAACCCAGGCGTACAGTGGGAGCT	1959
Db	1805	ctgactacacagagagcatccacgaagctgatacgaaactccagaaattgtctccagggagag	1864
QY	1960	ATATGAGAGCTGAGAAAGCAGGCGCCCGCTTCTGTGCCCCACTCTCCCAAGAACAGGTCC	2019
Db	1865	aggaagtcaagctcgaacgggtgtgtgaaagccagtcacagcccaagctgtgcccctgaagggtgco	1924
QY	2020	CCCATTTAGCCAGAGCTGACATTTGTTACCTTCAGAGGCGTCAACCGGGGCAACACCTTAG	2079
Db	1925	cccaactatgacagagctgtcatagtgaaactccaaggaatgtgaacgaaggaaggaacaacatct	1984
QY	2080	CTGTGCTGCTCACTGCCCCAGGGGAGCATC--GGGAGTGGGCCCCCGACAGTGGATTTCC	2136


```

QY 1660 TCAGCAGGCTGAGGAGGCTGTGTGGAGAGAGCTGACGTTACCTCTGTGCTCG 1719
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1655 tggagaagctcttcgagagatcgtgatacgtgaaatgcagtcacgttccttcgca 1714
QY 1720 GGGAGACTTCTCMTCAACCAACCCGCCAGGTCCTAGAGACCCCGTACCAGAGC 1779
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1715 gctatctagcatgttccaat-----aaocgtctct 1747
QY 1780 CCGGCTCTGTGGAGATCCGCCACTCCGCTCCCTGTGTCCTCCCAATGGCTCTGCTTC 1839
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1748 catcaccatgtaacaagggtccaactgcattacgatacgtcatt----- 1794
QY 1840 TGTCTCCAAATCCAGCCCTACCCCTCTGTGCGCCACTTACGCCGCTCCCTCCAGGCGC 1899
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1795 -----cccttcgc 1804
QY 1900 CGGGCCCCCACCACCCGCTGTGGCCAAACCCACCAACCCAGGCTTACGTGGGACT 1959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1805 ctgactaccagagacatccatccagctgatacgaataacccagaatttgcctcagggag 1864
QY 1960 ATATGAGCTGTGAGAGCAGGCGCCGCTCTGTGCGCCACTTCCCGAGAACAGGCTC 2019
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1865 agagatcagctgcagcggtgtgtgtaagccagtcacagccagtgccctgagggtgc 1924
QY 2020 CCCATTATGCGAGGCTGACATTGTACCTGTGACGAGGCGTACCGGGGCAACACTATG 2079
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1925 cccactatgagagctgatacagtaagtaagagtgacagggagagcaacataact 1984
QY 2080 CTGTGCCCTGCACTGCCCCCAGGGGCGATC---GGGGATGGGCCCCCAGATGGATTTC 2136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1985 cagtgccgcgcgcacacatcgtgacctgcctcaagaagatgtgctgtgaggagttcc 2044
QY 2137 CTGATCTGCACTCGCTTCAAGAGAGCTTGGCGAGGCGCATTTGGGGAGGGCGACC 2196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2045 ccaggaactccttaacttcaagaagagctgaggagagagcttggagggttccatc 2104
QY 2197 TGTGTGAGGTGACAGCCCTTAAGATCTGTCACTTGTGATTTCCCTTAATGTGCTA 2256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2105 tctgtgaagtgtgaggaaatcaaaatcaagaagaatttccctgagtgcagtg 2164
QY 2257 AGGAGACCCCTTGTGTGCTGACATCAAGATCTTACGGCCAGATGCCCAAGATGCCA 2316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2165 ccaaccagcgcctcctggtgctgtgaaatgcctcgcagatgacacagaagt--- 2221
QY 2317 GCTTCTCTGTTCTCCAGAGATTTCTGAAAGAGGTGAATCATGTGAGAGCTCA 2376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2222 -----ccaagaatgatttcttaaggagataaagatctcgtccca 2266
QY 2377 AGGACCCCAATATTTGCTGCTGTGGGCTGTGTGTGACAGAGACCCCTGTGACATGA 2436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2267 aggaccacaacatcatcatattatctgtgtatcacgtacgacctctcgtatga 2326
QY 2437 TTAAGTACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2327 tcaactgatacagagaaatgagatctcaatcagttcttccgcagacagaccctca 2386
QY 2497 ACAAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2387 attctctctcagcg-----atgtaagcactctcagtt 2419
QY 2557 ACCCAATGCTGCTGCATGTGGCAGCCAGATGCGCTCCGCGATGCTATGTGCCACAC 2616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2420 acacacatcgaagtattatgctacccaattgcctctgaaatgaatlaacttccctc 2479
QY 2617 TCACCTTGTACATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2480 ttaatttgttccagagatctgcccacagaaactgtttagtgtaaaataacacaa 2539
QY 2677 TCAAAATCGAGATTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2736
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2540 tcaagatagctgacttgagatgagcagaaactgtcagatgtgtgatacttaacggtacc 2599
QY 2737 AGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2796

```

```

Db 2600 agggccggcagctgtccctaccctcgtgagtgtcttggagagatcttgcctggcaggt 2659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2797 TCAGCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2660 tcaactacagcaagtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2719
QY 2857 GTAGGGCCAGCCCTTTGGGCACTCACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2720 tcaagaaacagccattccacagctgtcagatgaaagttatgaaatacttggaggt 2779
QY 2917 TCTTCGGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2976
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2780 tcttcgaagaccaaaggagagagacttccctcccaacccaattgtctctgactgt 2839
QY 2977 TATATGAGCTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2840 tglataagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2899
QY 3037 AGCTGATCGGTTCTT 3052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2900 aaatccactctgtct 2915

RESULT 11
AAT93784
ID AAT93784 standard; cDNA; 3157 BP.
AC AAT93784;
XX
AC AAT93784;
XX
DE 16-FEB-1998 (first entry)
XX
DE CCK-2, a human mammary carcinoma kinase 10 MCK-10) family member, cDNA.
XX
KW Mammary carcinoma kinase; MCK-10; CCK-2; receptor tyrosine kinase;
XX proliferative disease; cancer; MCK-10 activity; aberrant expression; ds.
XX Homo sapiens.
XX
OS
XX
FH Key
FT CDS Location/Qualifiers
FT 370..2934
FT /*tag= a
XX
PN US5677144-A.
XX
PD 14-OCT-1997.
XX
PF 08-NOV-1994; 94US-0336343.
XX
PR 16-NOV-1993; 93US-0153397.
XX
PA (ALVE/) ALVES F H E.
XX (ULR/) ULRICH A.
XX
PI Alves FHE, Ullrich A;
XX
DR WPI: 1997-511869/47.
DR P-PSDB: AAW3467L.
XX
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis
XX
PS Claim 9; Fig 3; 70pp; English.
XX
XX

```

The present sequence represents the cDNA sequence of human CCK-2, a member of the mammary carcinoma kinase 10 (MCK-10, AAT93785) family of receptor tyrosine kinases. Expression of CCK-2 is associated with proliferative diseases such as cancer. The CCK-2 gene was identified by PCR and a cDNA prepared from colon adenocarcinoma RNA. CCK-2 is expressed in a wide variety of cancer cell lines and tumour tissue. The CCK-2 nucleic acids can be used for diagnostic purposes to detect aberrant expression of CCK-2 genes. Engineered cell lines containing recombinant vectors with the present sequence, are useful for producing

CC Infectious retroviral particles. The cell lines may also be used to
XX evaluate and screen drugs involved in CCK-2 activation and regulation.
SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T; 0 other;

Query Match 16.2%; Score 642; DB 18; Length 3157;
Best Local Similarity 56.5%; Pred. No. 1,1e-130;
Matches 1534; Conservative 0; Mismatches 975; Indels 207; Gaps 10;

QY 349 TGCCTGCTCTGTTGGGCAAGTGGAGATGCTGACATGAAGGACATTTGATCTTCCCA 408
DB 395 TGGTCTGCTCTGTTGGGCAAGTGGAGATGCTGACATGAAGGACATTTGATCTTCCCA 454
QY 409 AGTCCGCTTATGCGGATGAGAGACCGGACATCCAGACAGATGACATCTCTCTT 468
DB 455 TAGCGCTATCTCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 514
QY 469 CCAAGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAGAG 528
DB 515 CCAAGTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAGAG 574
QY 529 ATGGGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 585
DB 575 ATGAGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 634
QY 586 TGAATACACAGAGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 645
DB 635 TGAATACACAGAGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 694
QY 646 GCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 705
DB 695 GCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 754
QY 706 GATGGCTGGAAGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 765
DB 755 GATGGCTGGAAGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 814
QY 766 GATGGCTGGAAGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 825
DB 815 GATGGCTGGAAGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 874
QY 826 GGGCTGACCGGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAG 885
DB 875 GGGCTGACCGGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAG 934
QY 886 ATGAGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAG 939
DB 935 ATGAGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAG 994
QY 940 CCGTGTACTCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 999
DB 995 CCGTGTACTCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1053
QY 1000 GTCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1059
DB 1054 GTCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1111
QY 1060 TGGGGGCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1119
DB 1112 TGGGGGCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171
QY 1120 ATGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179
DB 1172 ATGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1231
QY 1180 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
DB 1232 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1291
QY 1240 GTGGCCCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
DB 1292 GTGGCCCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1348

QY 1300 GGGACCCAGAGCCCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 1349 GGGACCCAGAGCCCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1408
QY 1360 AGTCCGCTTCTCTTCTGAG 1419
DB 1409 AGTCCGCTTCTCTTCTGAG 1468
QY 1420 ATGCTGGAACATTCCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
DB 1469 ATGCTGGAACATTCCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
QY 1480 CGCTGGCCACACCTCCACACATTCAGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1539
DB 1514 CGCTGGCCACACCTCCACACATTCAGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1574
QY 1540 CCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1599
DB 1535 CCGTGGGCAAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1654
QY 1600 TCTGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 1659
DB 1595 TCTGCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 1719
QY 1660 TCAAGAGAGAGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 1719
DB 1655 TCAAGAGAGAGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAG 1779
QY 1720 GGGACACTATCTCATGACACACAGAGATGCGGCTGATCCCGGAGATGCGGCTG 1779
DB 1715 GGGACACTATCTCATGACACACAGAGATGCGGCTGATCCCGGAGATGCGGCTG 1844
QY 1780 CCGGCTCTGCTGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1839
DB 1748 CCGGCTCTGCTGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1899
QY 1840 TGTCTCTCAATGACAGCTTACCGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1899
DB 1795 TGTCTCTCAATGACAGCTTACCGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1964
QY 1900 CCGGCTCTGCTGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1959
DB 1805 CCGGCTCTGCTGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 1984
QY 1960 ATATGAGCTGAG 2019
DB 1865 ATATGAGCTGAG 1924
QY 2020 CCGATTTATGCGGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 2079
DB 1925 CCGATTTATGCGGAGAGATTCACCGAGATGCGGCTGATCCCGGAGATGCGGCTG 2136
QY 2080 CCGTGTGCTGAGAGATTCACCGAGAGATGCGGCTGATCCCGGAGATGCGGCTG 2136
DB 1985 CCGTGTGCTGAGAGATTCACCGAGAGATGCGGCTGATCCCGGAGATGCGGCTG 2044
QY 2137 CCGTGTGCTGAGAGATTCACCGAGAGATGCGGCTGATCCCGGAGATGCGGCTG 2196
DB 2045 CCGTGTGCTGAGAGATTCACCGAGAGATGCGGCTGATCCCGGAGATGCGGCTG 2204
QY 2197 TGTGAG 2256
DB 2105 TGTGAG 2264
QY 2257 AGGAGACACCTTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAG 2316
DB 2165 AGGAGACACCTTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAG 2371
QY 2317 GCTTCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAGAG 2376
DB 2222 GCTTCTCTGCTGATGATTCATGCGCCGCGCACACAGAGATGAGAGAGAGAGAGAGAG 2466

```

QY 2377 AGAGCCCAACATCATTTGGCTGCTGGGCGTGTGTGTGACAGACACCCCTCTGCATGA 2436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2267 aggaccacaacatcatcatattatctgtgtatcaatgatacctctctgtatga 2326
QY 2437 TTACGACTACATGAGAAACGGACCTCAACACACTTCTCTAGTCCGCCACAGCTGGAGG 2496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2327 tcaatgaataatgtagatgagatccaatcagttcttctccgcagccccccta 2386
QY 2497 ACAAGGACCCGAGGGGCCCCCTGGGGAAGGGGACGCTGCGACAGGGCCACACACAGT 2556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2387 attctctctcagcg-----atgacacatgtcagtt 2419
QY 2557 ACCCAATGCTCTGCATGTGGCAGCCAGATCGCCTCCGGCAGTGCATGTGGCCACAC 2616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2420 acaccaatctgaagttaatgtagtgcacaaatgctctctgtacatgaactctctc 2479
QY 2617 TCACCTTTGTACATCGGAGACCTGGCCACGCGGAAGTGCCTAGTTGGGAAAATTTCACCA 2676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2480 ttaattgttccacagatctgtgcacacgaactgttttagtggtaagaactacacaa 2539
QY 2677 TCAAAATCGACACTTTGGCATGAGCCGGAACCTCTATGCTGGGGAGCTATTACGCTGTC 2736
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2540 tcaagataagctgactttggaatgagcaggaacctgtacagtgtgactatcacgagacc 2599
QY 2737 AGGCGCGGGCAGTGTGCTGCCATCCGCTGTGATGCGCTGGAGTGCATCCCTCATGCGGAAGT 2796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2600 agggccggcagtgctccctacatcgatgtctgtgagatgactctgtctgtgagagt 2659
QY 2797 TCACGACTGCGACAGTGTGGGCTTTGGTGTGACCTCTGTGGAGAGTGTGATGCTCT 2856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2660 tcaatgaagaagaatgtagtgtgagccttgggtacttctgtgagagacttcaacttt 2719
QY 2857 GTAGGGCCACGCTTTGGGAGCTCAACCGACAGAGCTATGAGAAACCGGGGAGT 2916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2720 gtcaagaacagccattatccagctgtcagatgacagagttatctgagatactgagagt 2779
QY 2917 TCTTCGGGACAGGCGCGGAGGTGTACCTGTCCCGGCTGCTGCTGCTGCGCAGGGCC 2976
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2780 tcttcagagacaaaggagagcagacttaacctccctcaacacagcattgtctgaactgt 2839
QY 2977 TATATGAGCTGATGCTTGGCTGTGAGCCGGAGAGTGTGAGCAGCAGCAGCCTTTTCC 3036
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2840 tgaatagcgtgagctcagctgtgagagagatacgaagaacgctccatccatcaag 2899
QY 3037 AGCTGATCGTGTCT 3052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2900 aaatccacctctcgt 2915

```

RESULT 12

AAVS5895
ID AAVS5895 standard; DNA; 3120 BP.

AC AAVS5895;

DT 19-NOV-1998 (first entry)

DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.

KM PTK; receptor; protein tyrosine kinase; brain tissue; ss.

OS Rattus sp.

FT Key Location/Qualifiers

FT CDS 485..3049 /tag="a" /product="PTK subtype tyro-10"

PN US811516-A.

PD 22-SEP-1998.

PF 02-JUN-1995; 9505-0456647.

```

XX 15-MAY-1992; 92US-0884486.
PR 02-MAY-1994; 94US-0237401.
PR 02-JUN-1995; 9505-0456647.
XX (SALK ) SALK INST BIOLOGICAL STUDIES.
PI Lai CHC, Lemke GE;
XX WPI; 1998-530939/45.
DR P-PSDB; AAW79152.
XX
XX Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably
XX expressed in brain tissue
XX
XX Example 2; Columns 53-60; 46pp; English.
XX
XX This DNA encodes a novel receptor protein tyrosine kinase (PTK)
XX polypeptide subtype tyro-10. The invention provides polynucleotide
XX sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13.
XX The PTK subtypes are found expressed predominantly in the brain tissue.
XX
XX Sequence 3120 BP; 760 A; 786 C; 830 G; 744 T; 0 other;

```

```

Query Match 16.1%; Score 639.8; DB 19; Length 3120;
Best Local Similarity 56.5%; Pred. No. 3,4e-130;
Matches 1552; Conservative 0; Mismatches 987; Indels 210; Gaps 11;
QY 348 CTGCTGCTGCTGTGTGGGCAAGTGGAGATGCTGCATGAAAGGACATTTGATCTCC 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 509 ctgtgtcgtctccgtcctctgtctacatcctgtgttgaataagctcagttatccagcc 568
QY 408 AAGTCCGCTATGCCCCCTGGGATGCGAGACCGACCATCCGACAGATGACATCTGCT 467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 569 atatgcgctatccctctgtgtatgtcagatgagagccaatccagatgagacacacagcc 628
QY 468 TCCAGCTCTGCTGATTCAGATTCACACTGCCCGCCACAGCAGATGGAGAGCAGTGCAGG 527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 629 tcaatcagtggtcagatccagatccagcgtgcacaaatattggtgagctgtcagaaaga 688
QY 528 GATGGGCGCTGTGCTGCCCGCAGAGGCTGCTTTCCCAAGA---GAGAGATCTTGCAG 584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 689 gatgagcctgtgtctcagatccagatccagatccagatccagatccagatccagatcc 748
QY 585 GTGATCTACACAGACCTCAGTGTGCTGCTGGGCGGACCGAGGAGGACGACGCGCG 644
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 749 atgactgtggaacccctacattatccatctctgtgtgagaccagggcgccatgacag 808
QY 645 GGCTGGGCAAGAGATTCTCCCGAGCTACCGGCTGCTTACCTCCGAGATGTGCGCGC 704
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 809 ggtatgcatlgtgaattgtcacccatgtacaaagatcaactacagtcgagatgagcagc 868
QY 705 TGGATGGGCTGGAAGAGACCCGCTGGGCTCAGAGAGTGTATAGCAATGAGACCTTGAG 764
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 869 tgaatccctcgtgtgtaacacgcatgagatgagatgagatgagatgagatgagatgag 928
QY 765 GGAATGTGTGTAAGAGACCTTGGGCGCCCGCATGATGAGGCGGATGCTGCTTACCC 824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 929 gatgtatccctcgaagactgtgagacacacatgtgtgcacagattgtgtcctatccca 988
QY 825 CGGCTGACCGGGGTGATGAGTGTGTGTGCTGGGCTAGAGCTATGCTGCTGAGG 884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 989 gtcaatgacacatccatgaaagtgtgcatgaggttgagcttatgtgtgtgtgtgcta 1048
QY 885 GATGAGCTCTGCTTACACCGCCCTGTGGGCGAGACATGATTTTATCTGAG----- 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1049 gatgtgtgtatctcctacatgctccagctgtgagagagattgtacccctgtgagctcc 1108
QY 940 -CCGTACTCTCAAGACACTCCACTATGACGACATACCTGGGCGGACTGAGTATGG 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1109 atcatattatctgataatctgtctatgatagtgagctgtgtgtgtacagcatgactgaagg 1168

```

999 GGTCTGGGCCAGCTGGAGATGGTGTGGGGCTGGATGACTTTAGGAAGAATCAGAG 1058
1169 ---ctaggccagttgactatgagttatccggccttgatgtattaccacccatgaa 1225
1059 CTGGGGGTCTGGCCAGGGATGACATATGTGGGATGGAGACCAACCAAGCTTCCAGTGGC 1118
1226 tacccagtgtagccgctgactgacagctggaatgagcgaagatgactacacagct 1285
1119 TATGTGAGATGAGTGTGAGTTGATGACGGGCTGAGGGGCTCCAGGATGAGCTCCAC 1178
1286 ttcattgagatcatgtttgaaatttgacggaatcagaattttacacatgaaagttccac 1345
1179 TGTAAACATGACACACCGCTGGGAGCCGCTGCTGGCGGGGAGAAATGTCGTTCCGG 1238
1346 tgcacaacatgtttgtctaaagtgtagaagattttaagagagtcagtgctacttcgc 1405
1239 GGTGGCCCTGCGCATGGCTGGAGAGGGAGGCCATGCCACAACTTAGGGGGCAACTG 1298
1406 tctg---aagccagagatggaacccactgctgctacttccctgtgcttgacgcat 1462
1299 GGGGACCCGACAGAGCCGGGCTGTCTCAGTGGCCCTTGGCGGGCGTGTGCTTCTG 1358
1463 gtgaaccccgagtgcccggtttgtacagtgccctccacacacagatgagccagtgcatc 1522
1359 CAGTCCCGCTTCTCTTGTGGGGCCCTGTGTTACTCTTCAGCGAAATCTCTCATCTCT 1418
1523 aagttccataaccatttgcgcgacagctgagatgttcagagagatcaacttccacaa 1582
1419 GATGTGAGAACATTTCTCTCCGGCAGTGGGAGGACCTTCCGGCCAGCCCCGGGTGG 1478
1583 gattgtgcaatgtataaacaactctgagcccttccacacttccca----- 1628
1479 CGGCTGGCCCACTCCACCAACTTCAGCAGCTTGGAGCTGGAGCCGAGAGCCAGAG 1538
1629 -----tggaaccacacacatgat 1648
1539 CCCGTGGCCAAAGCCGAGGGGAGCCGACCCGATCCATCTCATGTGCTGGTGGCCATC 1598
1649 cccatgtctaaagtgtatgatatgacaacactcgatccctgattgtgttgctgacatc 1708
1599 ATCTGCTCTCTCTCTCATATGCTCCCTCATGCTCTGTGGGGCTCAGTGGGGGAGGCTC 1658
1709 attctcaactgctgctgctatcctgctacatccctgtgagagcagttctgtgagaagatg 1768
1659 CTCACCAAGCTGGAAGGAGGTGTTGGAAGAGAGCTGACGGTTCACCTCTCTGCTCT 1718
1769 ctagaagaagcttcacgagagtgctgtagatgaaatgacacgacttcccttccctgccc 1828
1719 GGGGACACTATCTCATCAACAACCGCCAGGCTCTAGAGAGCCACCCCGTACCAAGAG 1778
1829 agcgagttccagatgttcaataacacacgctctc-----atacccaagtgaacagag 1882
1779 CCCCGGCGCTGAGGAAATCCGACCTCCGCTCCCTGTGTCCTCCCAATGAGGCTCTCGT 1838
1883 tccaacttactatgactcgaatcttcccttgccttgactacacagagacatccaga 1942
1839 CTGCTCTCCAAATCCAGCTTACCGCTCTCTTGGCCACTTACGCCCGTCCCTCTGAGGC 1898
1943 ctgact----- 1948
1899 CCGGGCCCCCAGACCCGCTGGGCAAAACCAACCAACCCAGGCTTACAGTGGGAGC 1958
1949 -----cgaatacttccagagttgtctccagagagagaga 1983
1959 TATATGAGCCTGAGAGGAGCGCCCGCTTGTGCCCCACCTCCCGAGAGAGCGGC 2018
1984 gtcaaggtgtagtggtgtgaaagcggcc-----cagcccaatggaactgaaagcgag 2038
2019 CCCCATATGCGGAGGCTGACATTTGTACCTGTCAAGGGCGTCAACCGGGGCAACACTAT 2078
2039 ccccaactgtagaagcagacatagtaatcccaagagagtagcaagtgtagcaaccctac 2098
2079 GCTGTGCTGCACTGCTGCCCGCAG---GGGCAGTGGGGGATGGGCCCCCAGAGAGTGAATTC 2135

2099 tgtgtgctgctgtaaccatgatctgcatcggggaagatgtgtcgtgaaagatctc 2158
2136 CCTCATCTCAGCTCCGCTTCAAGAGAAAGCTTGGCGAGGGGACATTTGGGGAGTGCAC 2195
2159 cccagaaactgtgtgctctcaagaagagcgtggaagagcgaatttggggaagtctac 2218
2196 CTGTGTGAGGTGACAGCCCTCAAGATCGGTCAGCTCTTGAATTTCCCTTAATGTGGCT 2255
2219 ctctgtgagatgagagagatggaataatccaagaacaaagatlttgactagatgctg 2278
2256 AAGGACACCCCTTGTGCTGAGCTGTCAAGATCTTAGCGCCAGTCCACCAAGATGCC 2315
2279 gccaacccagctgctgtgtgctgagaaatgctcgcagacagatgccaacaagatg-- 2336
2316 AGCTTCTCCTTGTCTCCAGGAATATTTCCGAAAGAGGTAAATATCATGTGAGGCTC 2375
2337 -----ccaagatgatttcttaagagagatcaagatcatgtctcgctc 2380
2376 AAGGACCCCAACATCATTTGCGTGGCGGTGTGTGTGACAGAGACCCCTCTGCATG 2435
2381 aagagcccaacatcatctcgtctcttaagctgtgtgcatcactgaaagacccgctgcatg 2440
2436 ATTACTGACTACATGAGAGAGCGGACCTCAACAGTTCTCAGTGTGCCACAGCTGGAG 2495
2441 atcaggaatacatgtgagaatgagatcttaactagttcttctgcacagagccctgt 2500
2496 GACAAGGACGCGAGAGGGGGCCCTGGGAGCGGAGGCTGCCCAAGGGGCCCACTCAGC 2555
2501 a-----gtctcttctcagtatgacagacagat 2530
2536 TACCAATGCTGCTCATGTGAGCAGCCAGATGCTCCGCGATCGGTATCTGCGACA 2615
2531 taagccaactgaagtattatgcaacccagatgtcctctgtatgaaatcatttcgct 2590
2616 CTCACTTTGTACATCGGAGCTGCGCCACGCGGAAGCTGCTAGTGGGAAATTTCAAC 2675
2591 ctcaacttgtccacagagatctgagcacaagaaactgtttagtgggaagatattacac 2650
2676 ATCAAAATGCGAGACTTTGGCATGAGCCGGAACCTTATGCTGGAGACTATACCGTGTG 2735
2651 atcaagaatagctgatttggcatgagcagaacccgttaagtgatattaccacagatc 2710
2736 CAGGCGCGGAGTGTCTGCCATCCATCCGTGATGGCTGGAGTGCATCTCATGGGGAAG 2795
2711 cagggccggggtgtctccacttgcgtgagatccttggaagagatcttgcctgagca 2770
2796 TTCAAGACTGCGAGTGAAGTGTGGGCTTTGTGTGACCTGTGGAGGTGCTGATGCTC 2855
2771 ttcaacacagcagatgagatgtgtgagccttggggtgactctgtgggaacacttccctt 2830
2856 TGTAGGCGCCAGCCCTTGGGAGGCTCAACCGAGAGAGGTATGAGAGCGGGGGAG 2915
2831 tgcagagagccactatcccaactgctgagtgagcagagttatcgaagaacctggagag 2890
2916 TTCTTCCGGAGACAGGCGCGAGGTGATCACTGTCCGCGCCCTGCTGCCGAGGCGC 2975
2891 ttcttcagagacaaagagagagatctatctccctcaacacacgcttggcccgact 2950
2976 CTATATGAGCTGATCTTCTGCTGAGACCGGAGTCTGAGACAGGACACCTTTTTC 3035
2951 gtgtataagctgtagtctcagctgtcgtgagaagaacaaacacgagccacttccag 3010
3036 CAGCTGATCGGTTCTCTGAGAGAGATGACATCAACACAGGTGTGAATCA 3084
3011 gaaataaccctctcttcttccagcaagagcagagatgattgactcatca 3059

RESULT 13

AAV65317

ID AAV65317 standard; DNA: 3120 BP.

XX AC AAV65317;

Oy 1049 GAGTCAGGAGCTGCGGGTCTGG 1070
|||||
Db 22 GAGTCAGGAGCTGCGGGTCTGG 1

Search completed: October 5, 2002, 20:29:15
Job time: 19933 sec